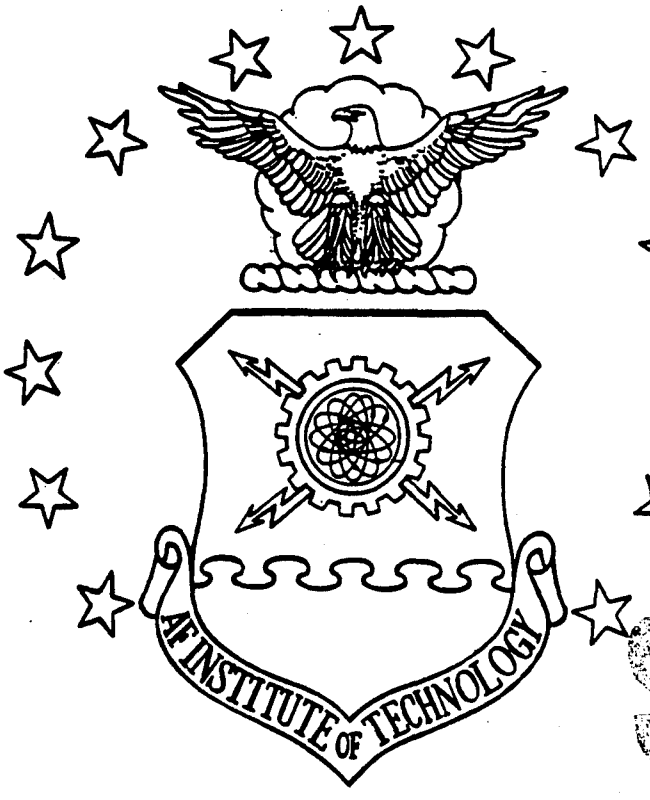


ead Related Transfer Function Approximation Using Neural Netv

THESIS

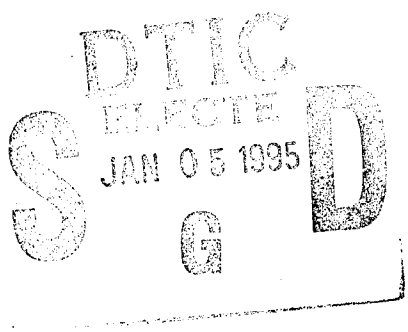
John K. Millhouse
Captain, USAF

AFIT/GE/ENG/94D-21

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GE/ENG/94D-21



Head Related Transfer Function Approximation Using Neural Networks

THESIS

John K. Millhouse
Captain, USAF

AFIT/GE/ENG/94D-21

Approved for public release; distribution unlimited

AFIT/GE/ENG/94D-21

Head Related Transfer Function Approximation Using Neural Networks

THESIS

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science (Electrical Engineering)

John K. Millhouse, B.S. Electrical Engineering
Captain, USAF

December, 1994

Acknowledgements

In thanking the many people who helped me, I will start at the very top. No one deserves more credit for this work than my Lord and Savior–Jesus Christ.

My thesis committee, composed of Dr. Steve Rogers, Dr. Mark Oxley, Barbara McQuiston, Dr. Dennis Ruck, and Dr. Matthew Kabrisky were the source of ideas, and technical expertise for the research in this thesis. A number of students were very helpful in teaching me better ways to use the Sun Workstations in the Signal Processing Lab. In particular Capt Brian Smith was extremely helpful with answering my questions regarding the use of ESPS and 3D sound in general. If not for the contributions of each of these individuals, my work load would have been even heavier. Finally, I would like to thank my wife, Audra, for her support and patience throughout this research effort.

John K. Millhouse

Table of Contents

Page

Acknowledgements ii

List of Figures vii

List of Tables xii

Abstract xiii

I. Introduction 1

1.1 Background 1

1.2 Problem 4

1.3 Definitions 4

1.4 Research Objectives 6

1.5 Scope 6

1.6 Approach 7

1.6.1 Synthesizing 3D Cues. 7

1.6.2 Artificial Neural Networks. 10

1.6.3 LNKmap. 13

1.7 Thesis Outline 14

II. Background 16

2.1 Introduction 16

2.2 Graphical Display 17

2.3 Graphical and Auditory Displays 18

2.4 Acuity of Sound Localization I 21

2.4.1 Interaural Time Difference 22

Page

2.4.2 Interaural Intensity Difference 22

2.4.3 Experiment 1 23

2.4.4 Experiment 2 24

2.4.5 Discussion 25

2.5 Acuity of Sound Localization II 25

2.5.1 Experiment 25

2.5.2 Results 25

2.6 Headphone Localization of Speech 26

2.7 Learning LNKmap with a simple function 27

2.7.1 Linear function 27

2.7.2 Non-linear functions 29

2.8 Learning LNKmap with a Cosine function 33

2.8.1 Cosine Revisited 34

2.8.2 Conclusion 36

2.9 Learning LNKmap with a Two Input Cosine function . . . 38

2.9.1 RBF Architecture 39

2.9.2 MLP Architecture 40

2.9.3 Conclusion 44

2.10 Conclusion 45

III. Methodology 46

3.1 Introduction 46

3.2 Learning HRTF Data for Zero Elevation 46

3.2.1 MLP Architecture 46

3.2.2 RBF Architecture 47

3.2.3 Conclusion 49

3.3 Experiment 1: With and Without HRTF 49

3.3.1 Objective 49

Page

3.3.2 Method 49

3.3.3 The Model 50

3.4 Experiment 2: AAMRL HRTF and RBF HRTF 54

3.4.1 Objective 54

3.4.2 Method 55

3.4.3 The Model 56

3.5 Conclusion 56

IV. Experiment Results 57

4.1 Introduction 57

4.2 Types of Error 57

4.2.1 Algebraic Error 57

4.2.2 Absolute Error 57

4.2.3 Reversals 57

4.2.4 Default to 90^o 58

4.3 Experiment One Results 58

4.3.1 Further examination of results 66

4.3.2 Summary of Experiment One Results 72

4.4 Experiment Two Results 72

4.4.1 Further examination of results 77

4.4.2 Summary of Experiment Two Results 81

V. Conclusions and Recommendations 84

5.1 Summary 84

5.2 Conclusions 84

5.3 Recommendations for Future Research 85

Appendix A. Experiment 1 Data: With and Without HRTF 86

Page

Appendix B. Experiment 2 Data: AAMRL HRTF and RBF HRTF . . . 101

Appendix C. Matlab M-files . . . 116

C.1 angleresponse.m . . . 116

C.2 anova.m . . . 117

C.3 drc.m . . . 121

C.4 expmat1.m . . . 122

C.5 expmat2.m . . . 123

C.6 idrc.m . . . 124

C.7 LRsphere.m . . . 125

C.8 makesphere.m . . . 125

C.9 mapsurf.m . . . 126

C.10 mapeval.m . . . 127

C.11 mypolar.m . . . 127

C.12 plotresult.m . . . 130

C.13 plothist.m . . . 131

C.14 readall.m . . . 133

C.15 removerev.m . . . 135

C.16 sortdata.m . . . 136

Bibliography . . . 137

List of Figures

Figure		Page
1.	Typical presentation forms of data	1
2.	Block diagram of a typical signal processing transfer function	7
3.	HRTF transforms sound to aid the brain in location perception	7
4.	Elevation, ϕ measured from the ground plane to a sound source	8
5.	Azimuth, θ of sound source to directly in front of face	9
6.	Distant, d of sound source to center of head	9
7.	Geodesic sphere with sound sources at multiple locations	10
8.	Multilayer-Perceptron (MLP) neural network with two hidden layers 10 nodes each	11
9.	Radial-Basis-Function (RBF) neural network	12
10.	Training function $y = \cos(2\pi x)$	13
11.	Training function $z = \cos(2\pi(x^2 - y^2))$	14
12.	272 speaker locations around the head	16
13.	HRTF samples from 0 to 1000	17
14.	Left ear HRTF data around the head for a frequency of 5623 Hz.	18
15.	Left ear HRTF data around the left side of head for frequencies 119 Hz to 19,953 Hz	19
16.	Zero elevation HRTF data for the left ear. Frequency and Response shown in logarithm scale	20
17.	A 3 layer net, 1 input node, 1 hidden node and 1 output node notated 1:1:1.	28
18.	$y = x$ (solid) and MLP (1:1:1) output (dash) after 100 epochs	29
19.	$y = x$ (solid) and MLP (1:5:1) output (dash) after 1000 epochs	30
20.	A 4 layer net, 1 input node, 5 hidden nodes,1 hidden node and 1 output node notated 1:5:1:1.	30
21.	$y = x$ (solid) and MLP (1:10:1) output (dash) after 1000 epochs	31

Figure Page

22. $y = x$ (solid) and MLP (1:10:1) output (dash) after 300 epochs . . . 31

23. $y \approx -x^2$ (solid) and MLP (1:10:1) output (dash) 32

24. $x \approx -y^2$ (solid) and MLP (1:10:1) output (dash) 33

25. Training function $y = \cos(x)$, where x ranged for 0 to 2π with a step size of 0.01 for a total of 629 points 35

26. Dynamic Range Compression of $y = \cos(x)$ to values between 0 and 1 35

27. $y = \cos(x)$ (solid) and RBF (1:1:1) output (dash) with 1 center in the K-means clustering algorithm 36

28. $y = \cos(x)$ (solid) and RBF (1:2:1) output (dash) with 2 centers in the K-means clustering algorithm 37

29. $y = \cos(x)$ (solid) and RBF (1:9:1, 1:32:1) output (dash) with 9 and 32 centers in the K-means clustering algorithm 37

30. $y = \cos(x)$ (solid) and MLP (1:10:10:1) output (dash) after 15, 100 and 220 epochs 38

31. $z = \cos(2\pi(x^2 - y^2))$ where x and y ranged from 0 to 1. 38

32. $z = \cos(2\pi(x^2 - y^2))$ where x and y ranged from -1 to 1. 39

33. Output of RBF with 8, 64 and 128 centers in the K-means clustering algorithm 40

34. RMS error of RBF with 8, 64 and 128 centers in the K-means clustering algorithm 40

35. Difference between outputs (left) and RMS errors (right) using 64 and 128 centers in the K-means clustering algorithm 41

36. Symmetric sigmoid used for non-linearity function $y = \frac{2}{1+e^x} - 1$ . . . 41

37. Output of MLP after 220 and 440 epochs with architecture of 2:10:1 42

38. RMS error of MLP after 220 and 440 epochs with architecture of 2:10:1 42

39. Output of MLP after 20, 50 and 100 epochs with architecture of 2:64:1 43

40. RMS error of MLP after 20, 50 and 100 epochs with architecture of 2:64:1 43

41. Output of MLP after 20, 220 and 440 epochs with architecture of 2:10:10:1 44

Figure Page

42. RMS error of MLP after 20, 220 and 440 epochs with architecture of 2:10:10:1 44

43. Zero elevation HRTF data for the left ear. Frequency and Response shown in logarithm scale 46

44. Zero elevation MLP (2:50:50:1) after 7235 epochs. Frequency and Response shown in logarithm scale 47

45. Zero elevation RBF with 256 centers. Frequency and Response shown in logarithm scale 48

46. Zero elevation RMS err for the MLP (left) and RBF (right) 48

47. Display provided on a computer screen for selection of the perceived location 50

48. Algebraic Error, positive values indicate a clockwise difference and negative values indicate a counter-clockwise difference 58

49. Reversal, a symmetric flip about the horizontal axis through the ears 59

50. Example of a subject with back-front reversals with HRTF treatment (left) and a different subject with front-back reversals without HRTF treatment (right). Notation: $\times$ indicates actual location, $\circ$ indicate perceived location. The line between the $\times$ and $\circ$ shows the pairwise groupings. The gradual spacing toward the center of the head is for clarity only and is not an indication of perceived distance. 59

51. Example of subject whose perceived locations defaulted to 90^o and 270^o. Notation same as above 60

52. Azimuth, θ of sound source to directly in front of face 61

53. Absolute mean error and standard deviation for sounds with the HRTF (solid) and without the HRTF (dash). The error bars show the standard deviation of the mean error. 67

54. Algebraic mean error and standard deviation for sounds with the HRTF (solid) and without the HRTF (dash). The error bars show the standard deviation of the mean error. 67

55. Mean and standard deviation for sounds with the HRTF (left) and without the HRTF (right). The arcs show the standard deviation of the mean (shown as a $*$). 68

Figure Page

56. Polar Histogram of all responses for sounds with the HRTF (solid) and without the HRTF (dash). 69

57. Polar Histogram of all responses correct within one button for sounds with the HRTF (solid) and without the HRTF (dash). 71

58. Polar Histogram of reversal corrected responses correct within one button for sounds with the HRTF (solid) and without the HRTF (dash). 71

59. Absolute mean error and standard deviation for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash). The error bars show the standard deviation of the mean error. 77

60. Algebraic mean error and standard deviation for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash). The error bars show the standard deviation of the mean error. 78

61. Mean and standard deviation for sounds with the AAMRL HRTF (left) and RBF HRTF (right). The arcs show the standard deviation of the mean (shown as a $*$). 79

62. Polar Histogram of all responses for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash). 80

63. Polar Histogram of all responses correct within one button for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash). 81

64. Polar Histogram of reversal corrected responses correct within one button for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash). 82

65. All subjects' responses for actual locations, 10^o, 32^o, 45^o, 58^o, 69^o and 82^o shown left to right, top to bottom respectively. Notation: $\times$ and solid line indicates with HRTF, $\circ$ and dashed line indicate without HRTF. The arcs show the standard deviation of the samples. The $\star$ and $+$ indicates the mean with and without HRTF respectively. . . 97

66. All subjects' responses for actual locations, 98^o, 111^o, 123^o, 135^o, 148^o and 169^o shown left to right, top to bottom respectively. Notation is same as above. 98

67. All subjects' responses for actual locations, 191^o, 212^o, 225^o, 237^o, 249^o and 262^o shown left to right, top to bottom respectively. Notation is same as above. 99

Figure Page

68. All subjects' responses for actual locations, 278^o, 291^o, 303^o, 315^o, 328^o and 349^o shown left to right, top to bottom respectively. Notation is same as above. 100

69. All subjects' responses for actual locations, 10^o, 32^o, 45^o, 58^o, 69^o and 82^o shown left to right, top to bottom respectively. Notation: $\times$ and solid line indicates with AAMRL HRTF, $\circ$ and dashed line indicate RBF HRTF. The arcs show the standard deviation of the samples. The $\star$ and $+$ indicates the mean AAMRL and RBF HRTF respectively. 112

70. All subjects' responses for actual locations, 98^o, 111^o, 123^o, 135^o, 148^o and 169^o shown left to right, top to bottom respectively. Notation is same as above. 113

71. All subjects' responses for actual locations, 191^o, 212^o, 225^o, 237^o, 249^o and 262^o shown left to right, top to bottom respectively. Notation is same as above. 114

72. All subjects' responses for actual locations, 278^o, 291^o, 303^o, 315^o, 328^o and 349^o shown left to right, top to bottom respectively. Notation is same as above. 115

List of Tables

Table		Page
1.	Simple linear function	28
2.	Simple nonlinear function	32
3.	Analysis of Variance for "Reduced three-way model"	55
4.	Experiment 1 factors	60
5.	Levels of "Azimuth"	61
6.	Analysis of Variance for "Response"	62
7.	Analysis of Variance for "Response" with reversal corrections	64
8.	Without reversal corrections (nor)	70
9.	With reversal corrections (rev)	70
10.	Summary of Experiment One	72
11.	Analysis of Variance for "Response"	73
12.	Analysis of Variance for "Response" with reversal corrections	75
13.	Without reversal corrections (nor)	80
14.	With reversal corrections (rev)	80
15.	Summary of Experiment Two	82
16.	Three dimensional graphic commands. Copyright (c) 1984-93 by The MathWorks, Inc.	116

Abstract

This thesis determines whether an artificial neural network (ANN) can approximate the Armstrong Aerospace Medical Research Laboratories (AAMRL) head related transfer functions (HRTF) data obtained from research at AAMRL during the fall of 1988. In order to test this hypothesis, two separate tests are performed.

The first test determines whether HRTF lends any support in sound localization when compared to no HRTF (Interaural Time Delay only). Depending on the statistical method used, we can conclude that the HRTF does provide an advantage in localization accuracy over simple ITD. There is, however, a statistically significant interaction between the location of the sound and whether the HRTF or no HRTF is used. When this interaction is removed using the alternate F-Value, the statistics give the result of equal means for the filters and azimuth. This means that at certain angles of azimuth, the HRTF either provides no advantage at all or hinders localization capabilities. Adding in the corrections for reversals changes the results to where the means of the azimuth are not statistically equal. The reversal corrections will inherently reduce the error results. However, comparing the number of reversals indicates an advantage of using the HRTF over no HRTF.

The second test determines whether HRTF and ANN lend the same amount of imformation in sound localization when compared to each other. Without considering reversals, we can conclude that the RBF HRTF provides a statistical advantage in localization accuracy over the AAMRL HRTF from which they were derived. However, with reversal corrections included, an advantage is not indicated, and two filters are statistically equal. Also with reversal corrections included, there is a statistically significant interaction between the location of the sound and whether the AAMRL HRTF or ANN HRTF is used. As discussed in the first experiment, this means that at certain angles of azimuth the HRTF either provides no advantage at

all or hinders localization capabilities. However comparing the number of reversals does not indicate a large difference between the AAMRL HRTF and the RBF HRTF.

The final conclusion is that the AAMRL HRTFs can be approximated by an artificial neural network for azimuth positions at zero degrees elevation and still have good results.

Head Related Transfer Function Approximation Using Neural Networks

I. Introduction

1.1 Background

The visualization of multivariate data is a complex area. There are a number of ways to help view the data. The conventional approach for the presentation of data is to present it in tabular or graphical form. Common graphical techniques work fine for three dimensions or less. For example, figure 1 shows some typical presentation forms of data. In each case, the presentation utilizes only the human sense of sight.

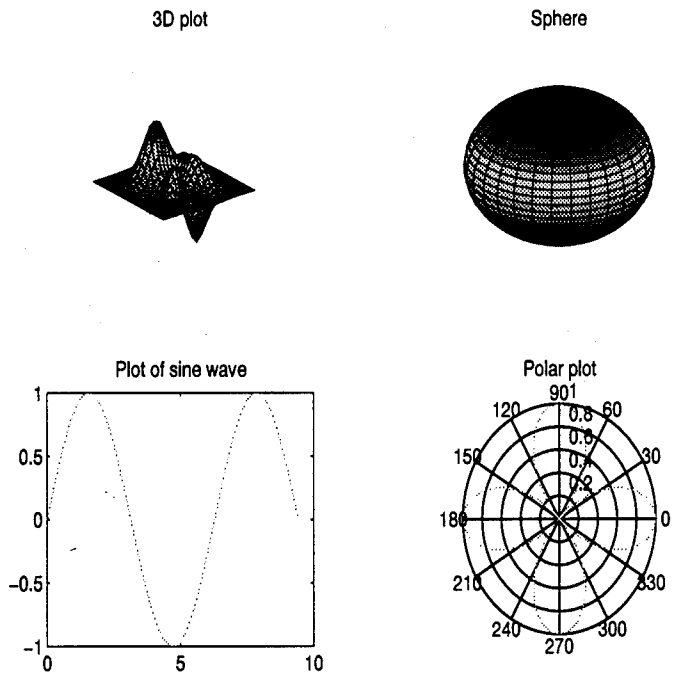


Figure 1. Typical presentation forms of data

The other four senses (auditory, touch, taste and smell) are not exploited. The

presentations in figure 1, as depicted, sound the same. Except for the thickness of the ink on the page, they feel the same. They also taste the same and smell the same. Of these four unused senses, the auditory sense has the most practical potential (6). For example, in a study of sleeping subjects, the subjects responded faster to an auditory alarm than to the presence of heat or the smell of smoke (10, 22). The fastest, simple reaction time come from electric shock (130 ms), followed by audio and tactile (140 ms) and finally visual (180 ms) (22). Sleeping is a case where a visual display is not appropriate. The following types of circumstances are possible cases where an auditory display would be preferable to a visual display: (22)

- When the origin of the signal is itself a sound
- When the message is simple and short
- When the message will not be referred to later
- When the message deals with events in time
- When warnings are sent or when the message calls for immediate action
- When continuously changing information of some type is presented, such as aircraft, radio range, or flight-path information
- When the visual system is overburdened
- When speech channels are fully employed (in which case auditory signals such as tones should be clearly detectable from the speech)
- When illumination limits use of vision
- When the receiver moves from one place to another
- When a verbal response is required

Many systems make use of different audio and visual displays. Sound displays are not limited to one dimension (monophonic) or two dimensions (stereophonic). Three dimensional (binaural) sound is where the listener hears the direction and

distance of the sound. Begault and Wenzel list some advantages and possible application of binaural systems: (3)

- monitor and identify sources of information from all possible locations
- improve intelligibility of sources in noise
- enhance segregation of multiple sources of speech
- improve air traffic control displays
- applications in cockpit communications
- telepresence applications such as teleconferencing
- shared electronic workspaces
- telerobotic control

A potential application listed above is in the cockpit (3, 8, 22). Although sound is already an important tool in the cockpit–only 1D sound is used. The pilot perceives the sound as coming from his or her earphones or from inside his or her head. In 3D sound, the pilot perceives the sound as coming not from the earphones or inside the head, but from a location outside the head at a certain azimuth, elevation and distance. By using 3D sound, the position of the pilot's wingman can then be encoded into three dimensions so that when the pilot hears the wingman speak he or she perceives the wingman's voice as coming from the wingman's relative position to that of the pilot. In the same way, targets and threats can be presented to the pilot. Active and non active threats could be distinguished by different pitches. The distance that the sound is perceived from the head could indicate the urgency of the data. Immediate action information could be placed close to the head while less urgent information could be placed farther away. Begault, in an experiment involving commercial pilots and collision warning systems, found that air crews presented with 3D auditory data acquired targets approximately 2.2 seconds faster than air crews that were presented only 1D auditory data (2).

In another example, data from an Army battlefield simulation was encoded into sound creating "battlefield songs" of the data (6). For example, in a song with only three variables: rhythm, tempo and pitch; rhythm may indicate the combatants of the battle, tempo may indicate the number of troops moving toward the front and pitch may indicate the number of troops already at the front. Using this scenario, a battlefield song of one rhythm that is increasing in tempo and pitch and another rhythm that has steady tempo and pitch would indicate one combatant of the battle had continuous reinforcements; while the other combatant did not get reinforcements and was losing troops at the front. In the multidimensional case, listeners had difficulty hearing the two sides of the battle independently; however, 3D sound may help separate the songs better by separating the songs to different locations outside the head (3).

1.2 Problem

This thesis determines whether an artificial neural network (ANN) can approximate the Armstrong Aerospace Medical Research Laboratories (AAMRL) head related transfer functions (HRTF) data obtained from research at AAMRL during the fall of 1988 (12). In order to test this hypothesis, two separate tests are performed. The first test determines whether HRTF lends any support in sound localization when compared to no HRTF (Interaural Time Delay only). The second test determines whether AAMRL HRTF and ANN HRTF lend the same amount of support in sound localization when compared to each other.

1.3 Definitions

This section provides definitions of key terms that will be used in this thesis. (25)

Binaural sound is sound that arises from two separate audio signals–one at each ear. In a natural environment, humans hear binaural sound. The signal that

arrives at the left ear is characteristically different from the signal at the right ear. The brain uses the differences between the two signals to determine the direction and distance of the source. Binaural sound is usually recorded with a manikin as opposed to stereo sounds which are recorded with without a manikin. The left ear signal is recorded separately from the the right ear signal. When sound is played through headphones, the sound is binaural if the signal sent to the left headphone is separate from the signal sent to the right (23:3).

Binaural mixing console is an electronic device that contains filters which convert a monophonic signal to a binaural signal corresponding to a given distance and direction (15:208).

Telepresence is a human's perception of being present in a natural environment while actually being present in an artificial or virtual environment.

Virtual audio is synthetically produced audio signals which enable a listener to achieve auditory telepresence.

Extracranialized sound is sound that is presented to a listener in the form of a binaural signal through headphones and is perceived as coming from some distance from the listener (i.e. outside the head).

Pinna(e) is(are) the human outer ear(s). The design of each person's pinnae is unique and each set of pinnae will uniquely filter sounds. In fact, the human head and ears form an antenna system characteristic to the individual (24). Experiments have shown that spectral shaping by the pinnae is dependent upon direction and cues provided by the pinnae. These are critical in extracranializing the sound (30:84).

Head related transfer function (HRTF) is the transfer function which accounts for the filtering effects of the pinnae. Because the filtering of sound by the pinnae is dependent upon direction, there is a different HRTF for each angle of azimuth and elevation. HRTF's will differ for each persons pinnae; however, these differences are

relatively small. Accurate localization of sound sources can be made using "someone else's pinnae" by presenting sounds filters with another persons HRTF's (31).

Auditory Localization Cues are cues that change the sound as a function of sound source and receiver location. There are four primary types of cues, 1) Monaural Temporal Cues, 2) Monaural Spectral Cues, 3) Binaural Temporal Cues, and 4) Binaural Spectral Cues. Spectral cues affect the spectrum of the sound that reaches the ear, while temporal cues affect only the arrival time of the sound signal (31)

1.4 Research Objectives

In order to solve the stated problem the following research objective will be met:

- Modify AFIT algorithms that have been used successfully on placing sounds in azimuth and elevation. The algorithms that create 3D sound with the AAMRL HRTF will be changed to use the ANN HRTF instead. This will allow the creation of tests to compare the ANN HRTF with the AAMRL HRTF.
- Statistically check the advantage of HRTF and ITD versus ITD only. The results of the first experiment will provide a significantly large data base to analyze.
- Statistically check the difference between AAMRL HRTF data and ANN HRTF data. The results of the second experiment will provide a significantly large data base to analyze.
- Investigate the program LNKmap as a neural network tool for function approximation.

1.5 Scope

This thesis will process and analyze experimental data to obtain results of the two experiments. The differences in sound localization accuracies from the HRTF

versus no HRTF experiment and the AAMRL HRTF versus ANN HRTF will both be analyzed by the analysis of variance method (14). To limit the amount of data to be processed only the 24 azimuth location at zero elevation will be analyzed. This will facilitate the Neural Network learning time and the amount of data to be analyzed.

1.6 Approach

1.6.1 Synthesizing 3D Cues. Humans use auditory localization cues to spatially localize aound. Three primary characterists of the auditory signals are attenuation, time of arrival (TOA) and HRTF. As the sound travels to the body, obstacles such as walls, furniture or even air attenuate the sound source. If the head is not directly facing the sound wave, the relative distance from each ear to the sound source will be different. This gives the brain the TOA cue. Once the sound reaches the body, the head, the torso, the nose and the outer ear or pinna filter the sound wave. This filtering function is known as the HRTF. The HRTF is analogous to an analog signal processing filter (figure 2). As in a analog filter where an input signal is

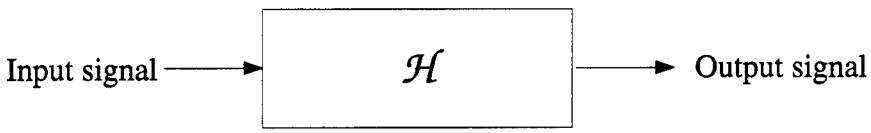


Figure 2. Block diagram of a typical signal processing transfer function

transformed by a system function, H to an output signal, the HRTF transformation of the sound appears to facilitate the brain in localization perception (figure 3). The factors which affect the HRTF include distance (d), elevation (ϕ), azimuth (θ),

Figure 3. HRTF transforms sound to aid the brain in location perception

frequency (ω), torso and head size. The frequency of the sound plays an important

role in determining the HRTF. Although the HRTF varies for frequency and for each point in space, the torso and head size can be generalized from person to person (3). For example, the HRTF for a sound located at 30^o elevation, 20^o azimuth, 10 feet and 440 Hz will be different from the HRTF for a sound located at 30^o elevation, 20^o azimuth, 10 feet and 1000 Hz. However, the sound transformed by the same HRTF will generally be perceived as coming from the same direction by people with varying size heads and torsos.

The elevation (ϕ) of the sound is measured from the ground plane through the ears to a sound source (figure 4). The azimuth (θ) of the sound is measured from

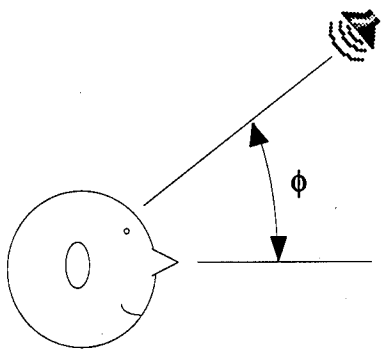


Figure 4. Elevation, ϕ measured from the ground plane to a sound source

directly in front of the face counter-clockwise to the sound source (figure 5). The distance (d) of the sound is measured from the center of the head to the sound source (figure 6).

All of these parameters can be measured in a sphere (figure 7) to determine the HRTF for any location and frequency. To measure the AAMRL HRTF, an anatomically correct manikin bust is positioned at the center of the anechoic chamber sphere with microphones placed inside the ear canals. Pure sine waves produced by the speakers are recorded with the ear microphones and used to determine the HRTF. The sine wave frequency is held constant until that particular HRTF is determined and then it is incremented for the next HRTF sample. The speaker location contains

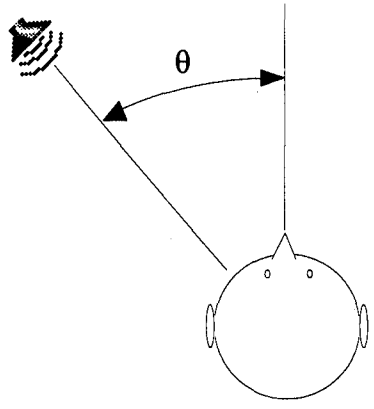


Figure 5. Azimuth, θ of sound source to directly in front of face

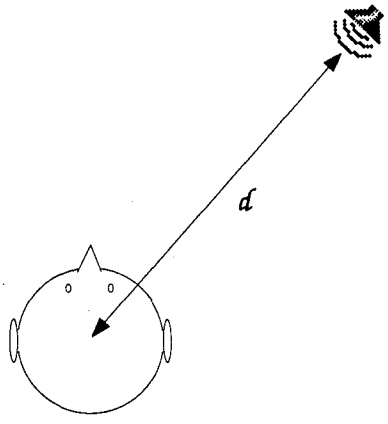


Figure 6. Distant, d of sound source to center of head

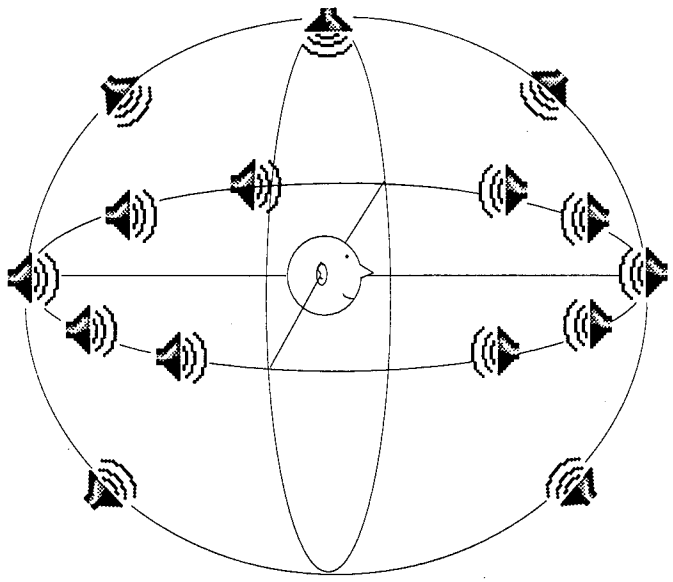


Figure 7. Geodesic sphere with sound sources at multiple locations (23:8)

the azimuth, elevation and distance information. Smith gives the location in azimuth and elevation of 272 speakers used to determine the HRTF by the above method (25). Once all the HRTFs are calculated, they can be used to filter any sound and its reflections. When attenuation and TOA are combined with the filtered sound, any recorded sound can be replayed to the listener through headphones to give the impression of a sound located at a distinct direction from outside the head.

1.6.2 Artificial Neural Networks. Current models of producing 3D sound from the HRTF require lengthy computations that cannot be accomplished in real time with a general purpose computer (26). Although two special purpose computers have been developed, the DIRAD and Convolvotron, this thesis will investigate the use of artificial neural networks to approximate the HRTF (23). Along with approximating the HRTFs, the networks will provide easy interpolation between the known HRTF data points. They may also provide insight into the underlying function of the HRTF. In other words, the networks may show that there is a less complex function for the HRTF than what is represented by the set of AAMRL HRTF data.

1.6.2.1 Multilayer Perceptron. Two types of artificial neural networks will be studied. The first is the Multilayer Perceptron (MLP). It has been shown by Cybenko that the MLP can approximate any arbitrary function (20). Figure 8 shows a three layer MLP (3:10:10:1). The input layer has three nodes corresponding to

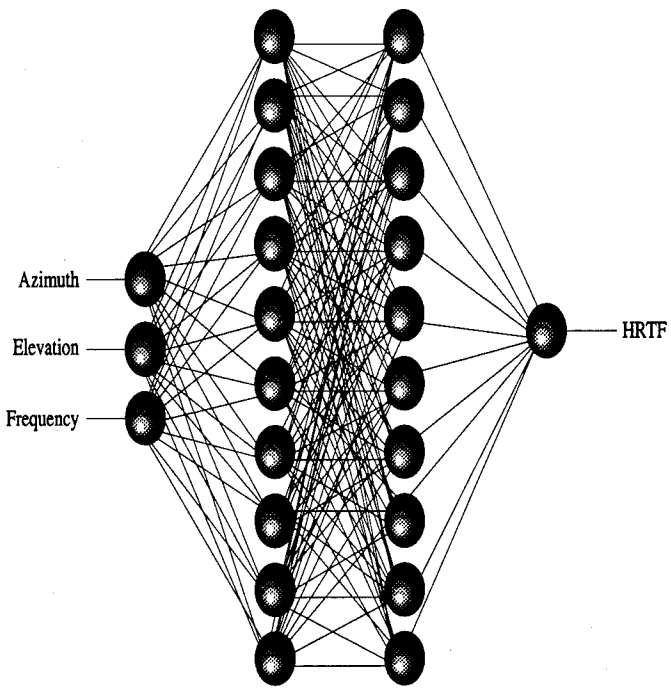


Figure 8. Multilayer-Perceptron (MLP) neural network with two hidden layers 10 nodes each

azimuth, elevation and frequency. The next two layers or hidden layers have 10 nodes each called hidden nodes. The last layer or output layer has one node corresponding to the HRTF. The number of nodes in the hidden layers are arbitrary, however, Neti, Young and Schneider found that a MLP network with less than 10 nodes in the hidden layers was sufficient for output direction from input sound (HRTF) (16). Although this thesis will reverse the inputs and outputs, input direction and output HRTF, it seems reasonable to start the investigation with a similar network architecture because it will basically be an inverse transformation of the Neti, Young and Schneider MLP.

The MLP will be trained with HRTF data from Smith's thesis (25). The data contain 25,296 HRTF taken from 272 speaker locations at 93 different frequency. The speaker locations range from 0^o to 360^o in azimuth and -90^o to 90^o in elevation. The frequencies ranges from 100 Hz to 20,000 Hz. A draw back of the MLP is the time required to train the network. Three layers of weights have to be trained. Therefore, the second network to be studied is the Radial-Basis-Function (RBF) artificial neural network.

1.6.2.2 Radial-Basis-Function. RBF networks are broad enough for universal approximation (19). The RBF is a two layer network. Figure 9 shows a RBF network with 8 nodes in the hidden layer. Again, the input layer has three

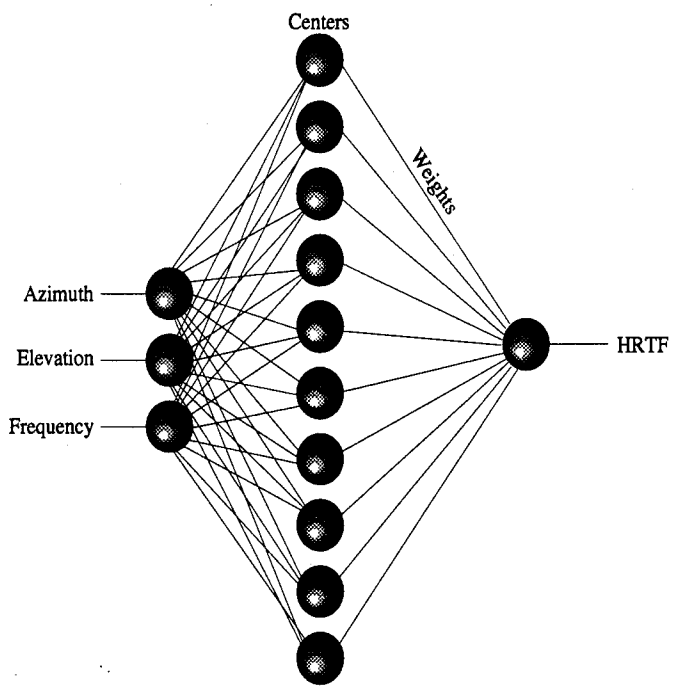


Figure 9. Radial-Basis-Function (RBF) neural network

nodes corresponding to azimuth, elevation and frequency. The next layer or hidden layer has 8 nodes corresponding to the centroids of the clustering algorithm. The last layer or output layer has one node corresponding to the HRTF output. The RBF will be trained with the same data as the MLP above.

1.6.3 LNKmap. The computer program LNKmap allows the user to train a number of different artificial neural networks on a SUN workstation. Both the MLP and RBF networks are available on this system. Since the HRTF is a complicated function to approximate, bugs in the LNKmap program would be hard to spot. Therefore, relatively simple functions were tested first to verify the inner workings of the LNKmap program. For example, a one input, one output function like $y = \cos(2\pi x)$ (figure 10) and a two input, one output function like $z = \cos(2\pi(x^2 - y^2))$ (figure 11) were trained and analyzed. From the training of these simple functions,

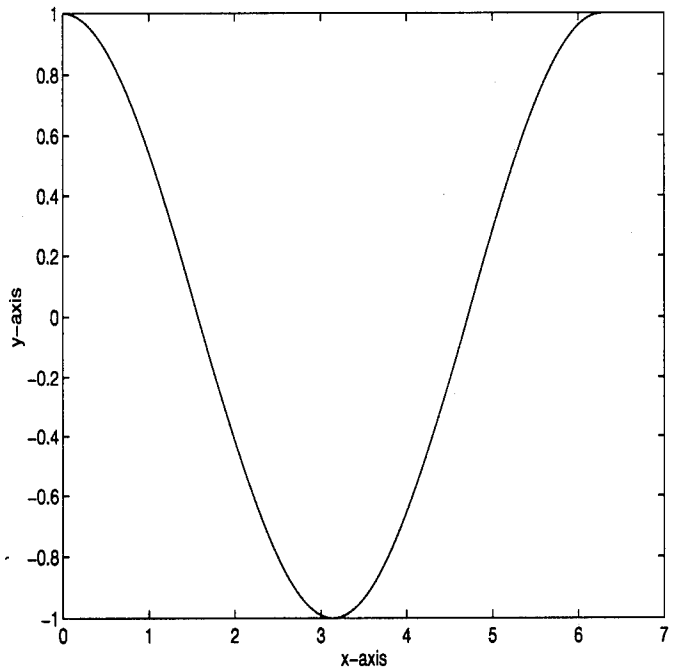


Figure 10. Training function $y = \cos(2\pi x)$

insight into the workings of LNKmap was gained. If LNKmap could not have been trained on one of these simple functions it would not have been able to be trained a complicated function like the HRTF. If this was the case, other programs would have been investigated to train the networks. Once a program was successfully trained on the simple functions, the process of training on the HRTF data was began.

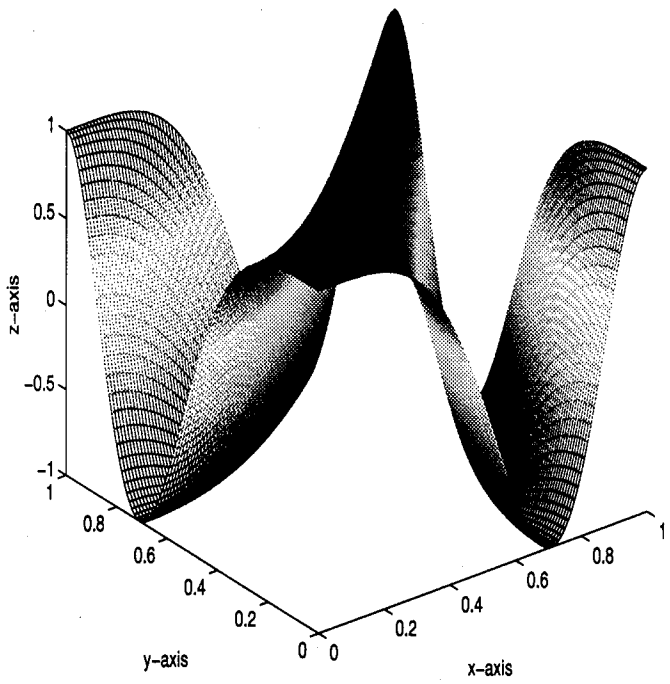


Figure 11. Training function $z = \cos(2\pi(x^2 - y^2))$

1.7 Thesis Outline

Chapter one is an introduction to this thesis. It presents a brief background of auditory displays which leads into the problem statement for this thesis. After that it defines key terms and research objectives. Finally it presents an approach to synthesizing 3D cues and creating articficial neural networks.

Chapter two continues the discussion of visualizing data that was introduced in chapter one. It also contains a review of past research in sound localization. Specifically, the work of Oldfield and Parker, and Begault and Wenzel is presented in great detail because of their direct correlation with this thesis. This chapter finally presents a historical record of the learning of the program LNKmap and the background needed to train an ANN with the HRTF data.

The previous chapters explored the background of 3D sound and ANN. Chapter three will combine these two concepts to train a ANN to approximate the 3D sound HRTF. This chapter presents the methodology used in training a MLP and RBF

ANN. It also presents the objectives and methodology of two experiments. Finally, it presents a detailed analysis of variance model used to analyze the experimental results.

Chapter four presents the results of the two experiments conducted during this thesis work. Before presenting the results of each experiment a discussion of the types of errors analyzed in the experiments is presented.

Chapter five presents the summary, conclusions and recommendations for future research.

II. Background

2.1 Introduction

The visualization of the data is important to help determine its characteristics. Viewing the HRTF data is a classic case of how to visualize multidimensional data. The HRTF data consists of 25,296 data points. Broken down, this consists of 272 azimuth and elevation locations: 0^o to 350^o in azimuth and -82^o to 82^o in elevation. The azimuths and elevations are spaced so that they are approximately 15^o apart over the whole sphere (see figure 12). For each azimuth and elevation pair, 93 frequencies

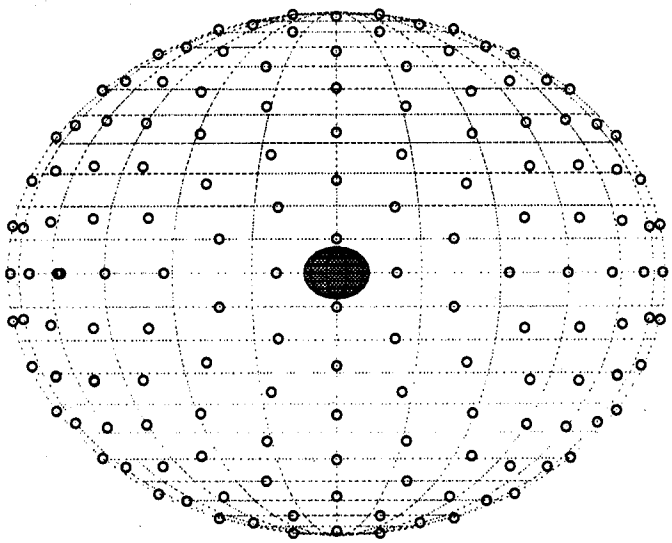


Figure 12. 272 speaker locations around the head

ranging from 100 to 19,952 Hz are used allowing the HRTF data to consist of four axes of data: azimuth, elevation, frequency and response. Figure 13 shows the first 1000 HRTF samples broken down into the four axes. Each line represents the value of the axis for that sample. For example, the 500^{th} sample approximately has a response of 0.05, an elevation of 70^o, an azimuth of 100^o and a frequency of 1000 Hz.

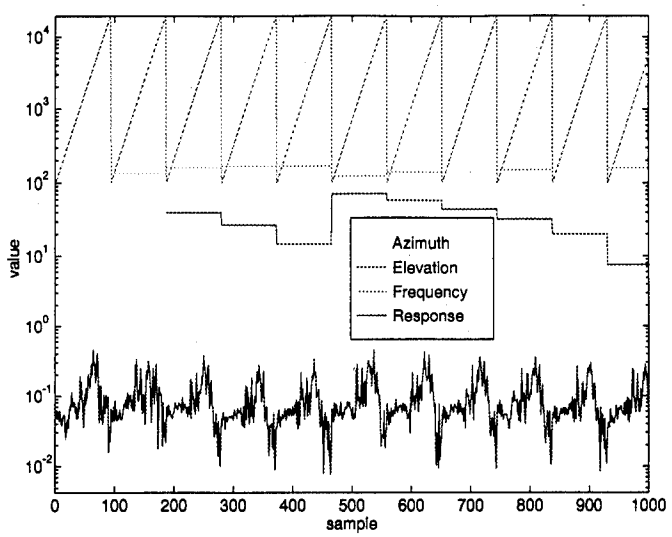


Figure 13. HRTF samples from 0 to 1000

2.2 Graphical Display

Matlab is a powerful tool to help visualize the data. One approach to visualizing 4 dimensions is to set the first three dimensions to the x, y, and z-axis and the 4th dimension to a set color. Matlab has a number of commands to do this.

By using the sphere command, a sphere around the head can be color coded with the HRTF value. The grey scale level indicates the magnitude of the HRTF. This display uses only the azimuth, elevation and response data. A separate sphere is needed for each frequency. Figure 14 shows left ear HRTF data around the head for a frequency of 5623 Hz. Note the lighter shades of grey on the left side of the head versus the right side of the head. This is because the left ear has a direct path from the sound source. Figure 15 shows most of the frequencies for the left side of the head only. The frequency increases from 119 Hz to 19,953 Hz. These two figures show graphically that the HRTF magnitude varies both as a function of location and frequency.

Similar to above where each plot consisted of a single frequency, figure 16 consists of a single elevation. The zero elevation data consists of 24 azimuth points

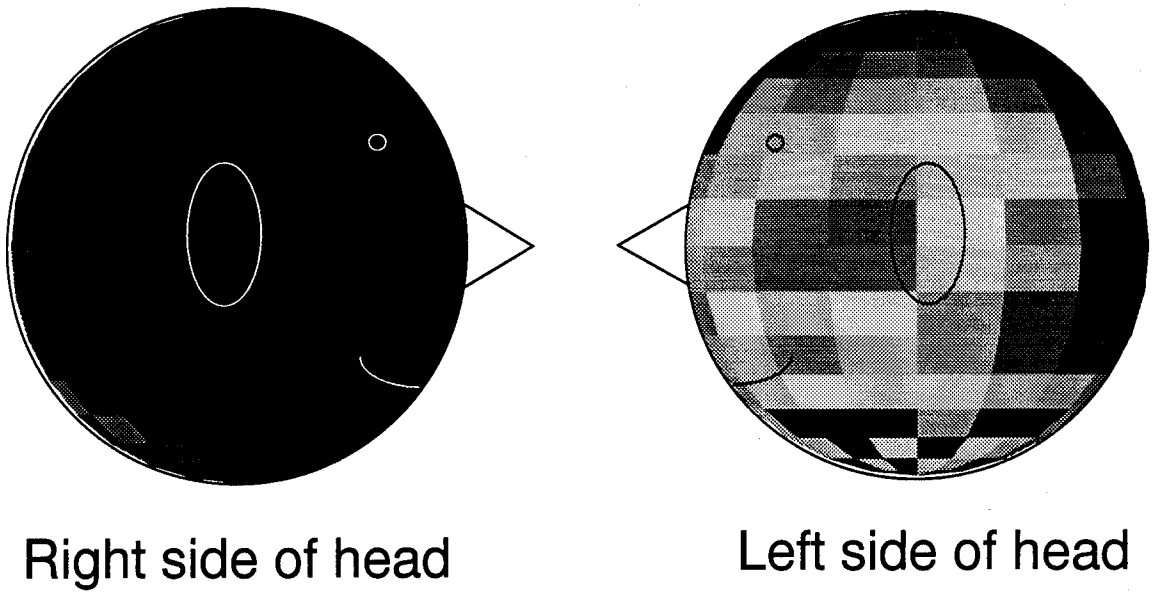


Figure 14. Left ear HRTF data around the head for a frequency of 5623 Hz.

with 93 frequencies each. For a total of 2232 data points. The response has a number of peaks and ridges with a high point localized slightly in front of the left ear at 3000 Hz. Both this view and the spherical view above have a limitations in the amount of data that can be presented graphically. The following sections present a discussion of the presentation of data.

2.3 Graphical and Auditory Displays

Graphical and auditory presentations are similar. When data varies with time, graphical presentation such as animation can provide a powerful analysis tool of graphical data (6). In the same sense, sound must be presented in time to be heard. If the length of the sound is too short its pitch, tempo, rhythm, and other attributes cannot be distinguished. Graphical symbols, known as icons, convey information in a small amount of space. On many computers with graphical interfaces, icons represent files, directories or functions. For example, a picture of a sheet of paper can represent a text document or a picture of a trash can represent a delete file function. "Earcons," the audio counterpart of icons, similarly convey information (4). The

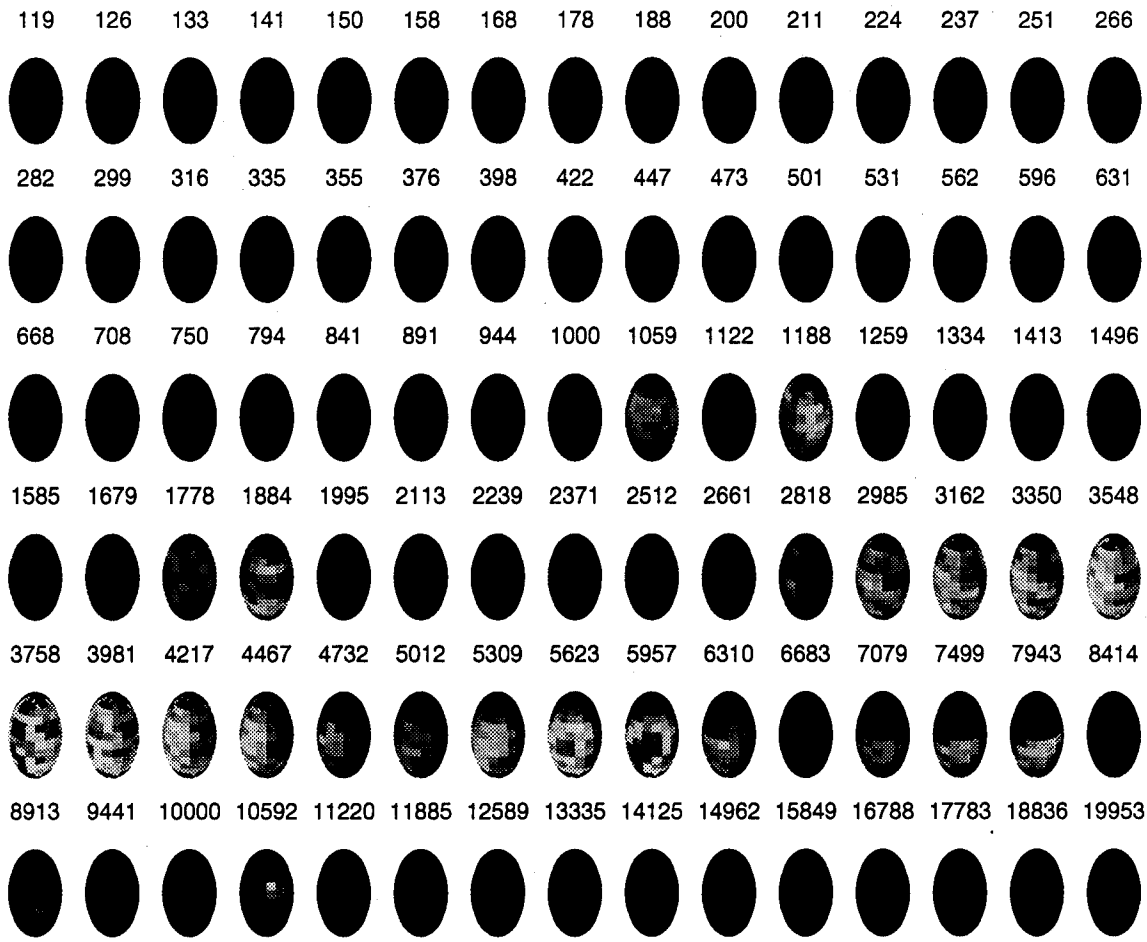


Figure 15. Left ear HRTF data around the left side of head for frequencies 119 Hz to 19,953 Hz

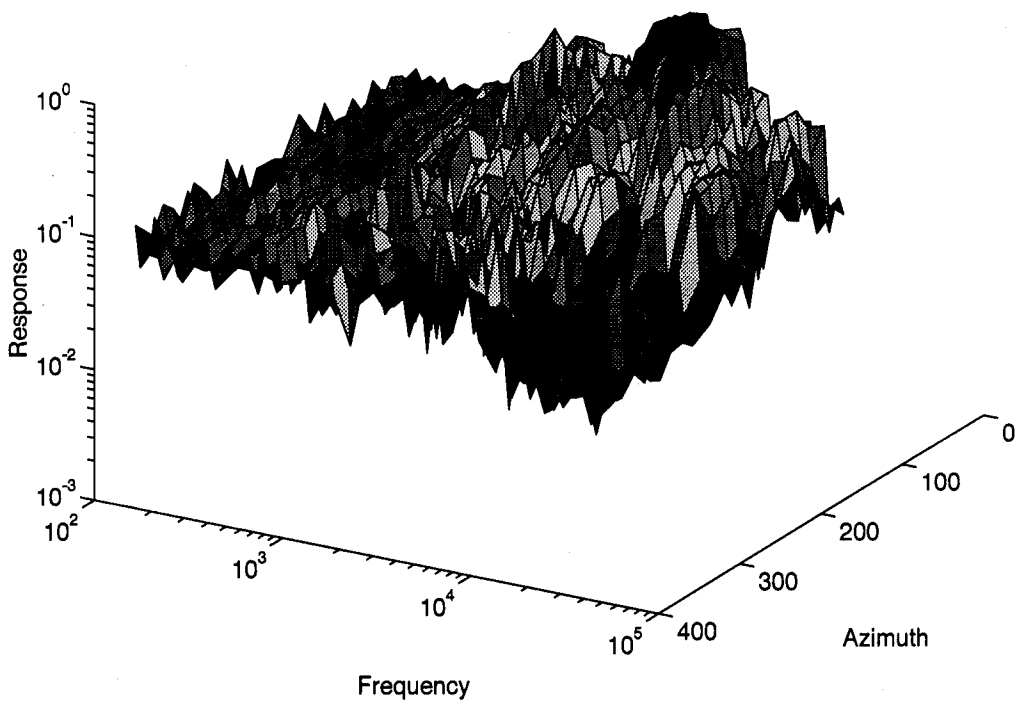


Figure 16. Zero elevation HRTF data for the left ear. Frequency and Response shown in logarithm scale

beep of the computer when it displays a warning message is an earcon. The two tones of a door bell is an earcon conveying the message that someone is at the door. In graphical presentation, color is often used as the fourth dimension when presenting 4D data. Sound resolution is similar to color resolution (6). Both the frequency of pitch and the frequency of color are logarithmic. (blue 435.8 nm, green 546.1 nm and red 700 nm) (9) Octaves provide expression of logarithmic variance (6). On the other hand, sound is different from a graphic display. As discussed above, the graphical presentation of multivariate data is commonly limited to three or four dimensions. In contrast, the human ear can discriminate between any two of 400,000 different sounds presented in rapid succession. In addition, most people can identify 49 different sounds at one time (6). The different sounds can be used to represent different dimensions of data. For example, pitch, volume, duration, attack, fundamental wave shape and fifth harmonic wave shape can be used to represent 6 different dimensions. Yeung used nine to twenty dimensions of sound (6). Matthew used a combination of auditory and visual presentations of data (6). Matthew had five dimensions, x and y visually and frequency, timbre and amplitude phonically. Smith, Bergeron and Grinstein also combined auditory and visual presentations of data (26). In their scheme, the auditory data representation is triggered by the position of the mouse on the visual display. Most of the past research has been with 1D and 2D sound. Although Bly did not use 3D sound in her thesis, she suggests using 3D sound for the visualization of three dimensional data (6). "Several sound characteristics were not used and these deserve attention in the future. Certainly location is an easily detectable attribute of sound." (6:40)

2.4 Acuity of Sound Localization I

In an effort to investigate the acuity of sound localization, Oldfield and Parker conducted a study of 8 subjects who judge the apparent spatial location of white noise through speakers over a range -40^o to $+40^o$ in elevation and 0^o to 180^o in

azimuth (17). There are two major binaural cues: Interaural Time Difference (ITD) and Interaural Intensity Difference (IID).

2.4.1 Interaural Time Difference.

- Interaural Time Difference is a function of the distance between the two ears, the angle of incidence of the incoming sound, and the frequency of the sound (25).
- ITD is frequency independent below approximately 500 Hz and above approximately 3000 Hz; furthermore, between these frequencies ITD decreases as frequency increases. For azimuth angles of incidence less than 60 degrees left or right of the front of the listener (0 degrees), the minimum ITD occurs between 1400 and 1600 Hz. (25, 11:165-166).
- The most significant changes in ITD occur at angles of incidence between 0 and 30 degrees, consistent with the fact that humans perceive a sound best when the source is directly in front of the listener (25, 1:700).

2.4.2 Interaural Intensity Difference.

- Interaural Intensity Difference is a function of the distance between the two ears, the angle of incidence of the incoming sound, and the frequency of the sound. IID can also be affected by size and shape of the torso, head, and ears. The torso affects primarily the low frequencies (25).
- The most significant changes in IID occur at angles of incidence between 0 and 30 degrees, consistent with the fact that humans perceive a sound best when the source is directly in front of the listener (25, 1:700).
- Improvement in localization capabilities of humans at frequencies above 3000 Hz is a result of improvement in the IID cue. (25, 11:166).
- At frequencies above 8000 Hz, IID cues increase human capabilities to localize in elevation as well as in azimuth (25, 13:101).

From these two cues, there is a range of positions which could account for the same pattern of interaural differences. In a simplified model of the head, these positions form a cone. Watkins calls it the "cone of confusion" (29). Oldfield and Parker presented a second set of cues which they describe as spectral modifications by the pinna and head. These spectral modifications are what Begault and Wenzel call the head related transfer function (HRTF) (3). Past studies show evidence of non-binaural cues (HRTFs) (17). These studies use sound sources in the median sagittal plane. The median sagittal plane divides the body into left and right sides. Other studies blocked the pinnae or low pass filtered the sound sources and confirmed that localization ability is lost. Few studies before 1984 use sound sources outside the median sagittal plane (17).

2.4.3 Experiment 1. The Oldfield and Parker experiment was conducted in an anechoic chamber for frequencies down to 250 Hz. The apparatus for the experiment consisted of a boom for positioning the sound source, a photographic recording system and pointing gun. The subject's head was held stationary. The importance of head motion in human auditory localization cannot be overstated. Without head motion, a person's head and ears can still act as a directional antenna; however, sounds located at 0 degrees azimuth and various elevations (i.e. on the medial sagittal plane) can be extremely difficult to localize. Front/back and up/down confusions result when head motion is not allowed; this occurs because the spectral and temporal cues are very similar at both ears (31, 5, 25). The implementation of head motion into a binaural room simulation requires that a real-time system be used (23). For this reason, head motion was not implemented into the non-real-time simulation done for this thesis.

The sound was white noise which the subject could activate through a hand-held switch. The subject used a pointing gun to "shoot" the sound. The position was recorded with the photographic recording system. The advantage of pointing is that the response is not limited to preset locations. The disadvantage of pointing is

the limitation of the accuracy with which the hand can be positioned. Fitts found that blind-positioning movements are most accurate in the dead-ahead, center and lower tier positions and least accurate at the side and upper tier positions (22, 7) To avoid pointing inaccuracies in this thesis, the sound locations were fixed and subjects were required to pick preset locations.

2.4.4 Experiment 2. To validate the pointing technique, Oldfield and Parker conducted a second experiment to the measure the motor skills of the subjects. The apparatus and procedures were the same as experiment 1 except the subjects were not blindfolded and the chin brace was loosely fitted. The subjects were allowed to see the speaker position before the sound was presented. Next the room was completely darken and the sound presented for the subject to "shoot." They concluded that the biases in the pointing response did not affect the overall outcome of the results. Reversals were never observed in the visual task (17).

2.4.4.1 Results. Azimuth error and elevation error were measured. The absolute and algebraic form of these were computed. Absolute error measured general acuity and algebraic error measured directional bias. For algebraic error, positions perceived higher and behind were termed positive errors (actual position minus perceived position). Front/back reversal were analyzed separately: "a straight subtraction of the actual and perceived sound position has not been regarded by researchers as an honest representation of error (17, 28:587)." This thesis also analyzed reversals.

There were few front/back reversals. Most occurred in the upper back quadrant. One exception to this was the region near 90^o azimuth where the reversals were uniform over elevation. Another kind of error was what Oldfield and Parker called "defaults to 90^o". Like reversals the defaults occurred in the upper back quadrant. Defaults tended to boost the error in the upper back quadrant, whereas reversals and non-reversals had the same amount of error.

2.4.5 Discussion. The results support the model of the cone-of-confusion in the case of elevation error biases effecting azimuth error biases. Both azimuth and elevation discrimination is less accurate behind the head. Oldfield and Parker suggest that sound comings from behind the head do not directionally interact with the convolutions of the pinna. Front/back discrimination derives from the pinna when the head is held still. The higher number of reversal in the back region is consistent with a reduction in pinna cues. Binaural differences cannot be separated from pinna modification because they are in the signal before any binaural differences can be processed by the brain.

2.5 Acuity of Sound Localization II

Oldfield and Parker also conducted a study of 4 subjects who judged the apparent spatial location of white noise through speakers over a range -40^o to $+40^o$ in elevation and 0^o to 180^o in azimuth (18). In this experiment the pinna of the subjects were blocked to simulate the effect of localization without the pinna. This experiment has a direct correlation with the first experiment conducted in this thesis where the hypothesis of no HRTF is similar to no pinna. The implication is that the determination of azimuth and elevation depends on more than just the binaural differences from the sound source. An additional cue is the spectral modification made by the pinna.

2.5.1 Experiment. The data collection apparatus was the same as described above in part I of the study (17). Each subject was fitted with individually cast molds for the pinna. An access hole remained to the auditory canal. Subjects were presented 171 different speaker locations while blindfolded and head help stationary.

2.5.2 Results. The Oldfield and Parker study presented the relationship between azimuth error and elevation error. This thesis presents only azimuth error. With the pinna blocked the overall absolute elevation error was almost double that

of azimuth error (11.9^o and 21.9^o respectively). The largest differences between absolute azimuth and elevation errors occurred in the front quadrant. The absolute azimuth error was largest between 120^o and 160^o azimuth position. Azimuth error is reasonably uniform and elevation error is less uniform with larger errors near the extremes. The mean algebraic elevation and mean azimuth errors were generally positive in the front and negative in the back. The subjects perceived the sounds in the front pulled higher and toward the ears. The subjects also perceived the sounds in the back pulled lower and toward the ears. Filling the convolutions of the pinna has the effect of displacing elevation judgements toward 0^o.

A comparison of the results of the pinna filled versus unfilled (normal) showed that there was a increase in elevation error in the pinna-filled case (filled 21.9^o versus normal 8.4^o). The azimuth error also increase in the pinna-filled case, but not as great (filled 11.9^o versus normal 9.3^o). The mean absolute azimuth error for the filled-pinna was generally lower than the normal pinna except to azimuth positions 140^o to 160^o. The error did not change significantly by azimuth position, however it did change significantly by elevation position.

A number of subjects made front/back reversals. Reversals accounted for 26% of the responses: 31% were front/back reversals and 21% were back/front reversals. These results are compared with the findings in this thesis in chapter 4.

2.6 Headphone Localization of Speech

A study of 11 inexperienced subjects who judge the apparent spatial location of headphone-presented speech stimuli filtered with non-individualized head-related transfer functions (HRTFs) was conducted by Begault and Wenzel. (3). This is in contrast with the Oldfield and Parker study who used experienced subjects listening to white noise from an actual sound source in three dimensional space.

HRTFs are "the listener-specific, direction-dependent acoustic effects imposed on an incoming signal by the pinnae (3:361-362)." It should be noted that not only

the pinnae, but the head and torso effect the incoming signal also. Note also that this definition says "listener-specific". Although HRTF are "listener-specific", they generalize to other people. The main thrust of the Begault and Wenzel study is to investigate how non-individualized HRTFs work. The HRTFs used are from the pinnae of a good localizer (or as Dr Rogers' would say a person with a "golden ear"). The study decides against simple averages of HRTFs to avoid possibly elimination of distinctive spectral features. This thesis used AAMRL HRTFs (12).

Eleven adults were used in the Begault and Wenzel study. They were screened by oral questions about hearing loss, recent exposure to loud noises and work environment. There is no apparent relation between audiometric measurements and binaural performance (3). Therefore this thesis will forego the audiometric measurement of subjects.

The sounds used in the study were from a set of 45 one- or two-syllable words representing a particular phoneme. The sounds were filtered with the HRTFs to simulate sounds coming from 0^o elevation and 0^o, $\pm 30^o$, $\pm 60^o$, $\pm 90^o$, $\pm 120^o$, $\pm 150^o$ and 180^o azimuth. Similarly, this thesis used the male voice utterance "seventeen" presented at 0^o elevation and 24 azimuth locations. Begault and Wenzel conclude that most listeners can obtain useful directional information from speech without requiring individual HRTFs. Taking this result a step further, this thesis investigated the hypothesis of obtaining useful directional information from speech filtered with a neural network trained HRTF.

2.7 Learning LNKmap with a simple function

2.7.1 Linear function. During the spring of 1994, the program LNKnet was gaining popularity at AFIT as a neural network classifying tool. LNKnet's companion program for function approximation (LNKmap) had not yet been fully investigated. To begin the exploration of LNKmap, a simple function i.e. $y = x$ was inputed into LNKmap.

Since LNKmap was similar to LNKnet the file structure was assumed to be the same. Therefore the input data file should be in two columns with the output values in the first column and the input values in the second column. A small file was used starting with 10 points between zero and one (see table 1). Since the

y	x
.1	.1
.2	.2
.3	.3
.4	.4
.5	.5
.6	.6
.7	.7
.8	.8
.9	.9
1.0	1.0

Table 1. Simple linear function

function is linear, a linear output node function was used. Most network users call the "output node function" a "non-linearity function". A common starting point is a 3 layer multi-layer perceptron (MLP), 1 input node, 1 hidden node and 1 output node notated 1:1:1 (figure 17). Some researchers refer to this architecture as a 2

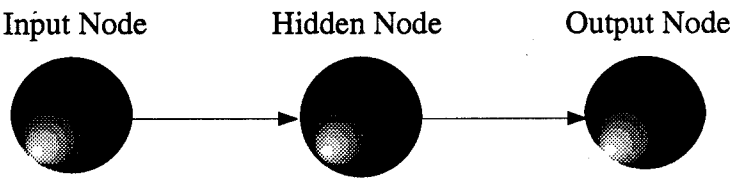


Figure 17. A 3 layer net, 1 input node, 1 hidden node and 1 output node notated 1:1:1.

layer net referring to the number of weight layers. This thesis will try to use the colon notation where ever a confusion may exist. Many classifiers present the data in random order, however for function approximation it is best to present the data in order (21). With the training set described above the net learned the function in about 100 epochs.

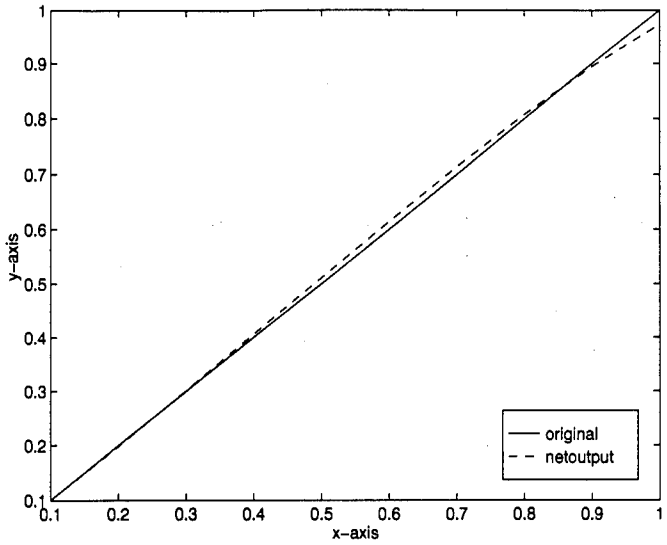


Figure 18. $y = x$ (solid) and MLP (1:1:1) output (dash) after 100 epochs

Next a standard sigmoid on the same data was tried while also increasing the number of hidden nodes to 5 (1:5:1). The net learned the function in about 1000 epochs (see figure 19). In an effort to get a lower error a second hidden layer with one node (1:5:1:1) was added. This architecture is shown in figure 20. After a 1000 epoch, the net had not learned the function. More nodes to the second hidden layer (1:5:5:1) were added. After another 1000 epoch, the net had not learned the function.

Adding more nodes seems to slow the learning process therefore one hidden layer was used in subsequent trials. With an architecture of 1:10:1 and a standard sigmoid output node function, the net learned the function (see figure 21). With the same architecture and using the symmetric sigmoid output node function, the net learned the function with an even lower error (see figure 22).

2.7.2 Non-linear functions. At this point, some nonlinear data was tried. The 10 point data was modified to be low near the end points and higher in the middle ($y \approx -x^2$, see table 2 and figure 23). The LNKmap plot of the data approximated the function $x \approx -y^2$ (see figure 24). This indicated that the columns of the data file were reversed. This discovery uncovered that LNKmap needed the

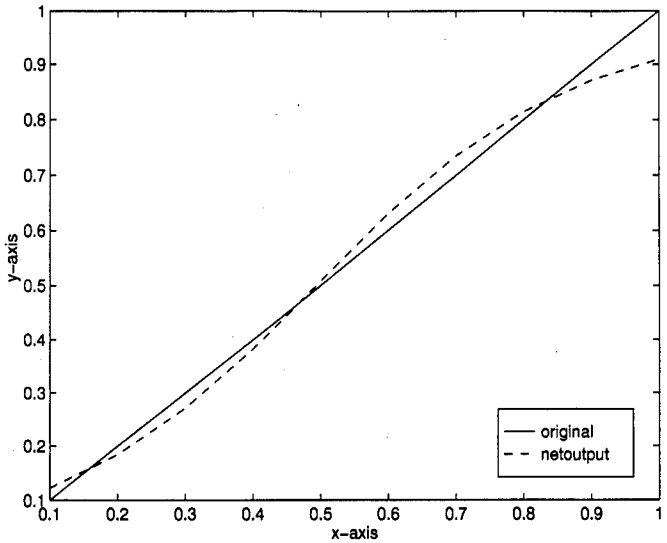


Figure 19. $y = x$ (solid) and MLP (1:5:1) output (dash) after 1000 epochs

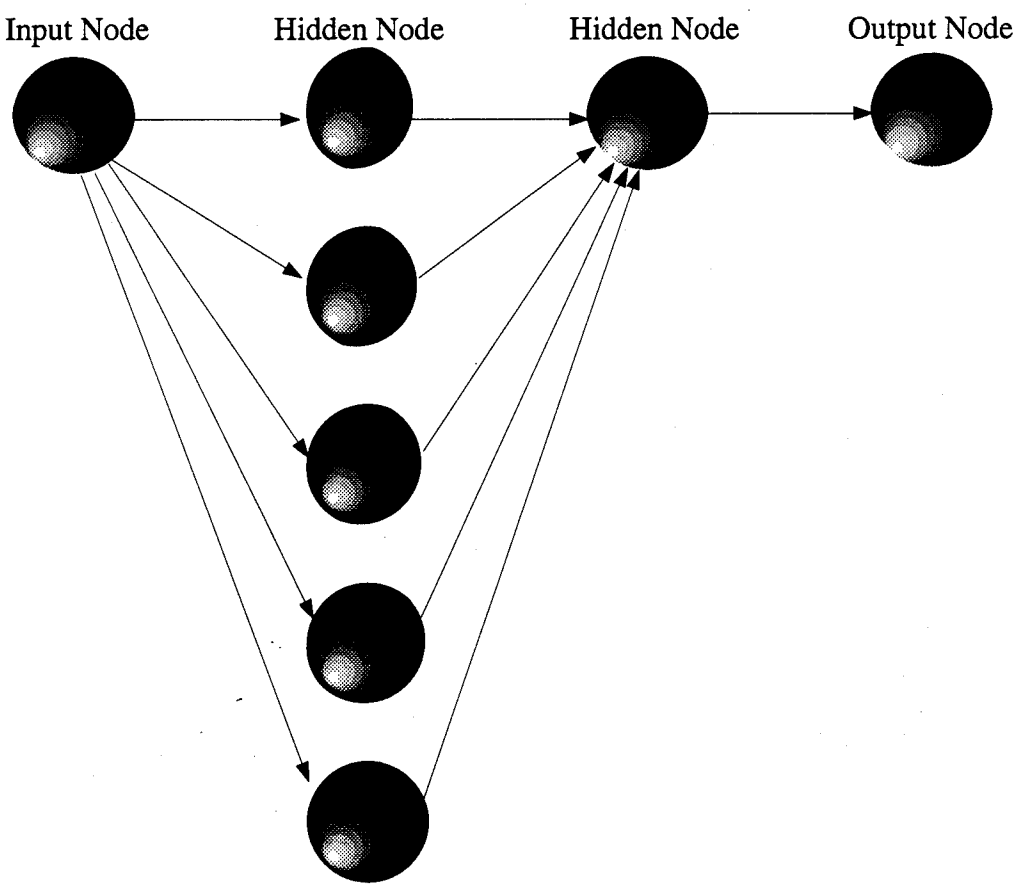


Figure 20. A 4 layer net, 1 input node, 5 hidden nodes,1 hidden node and 1 output node notated 1:5:1:1.

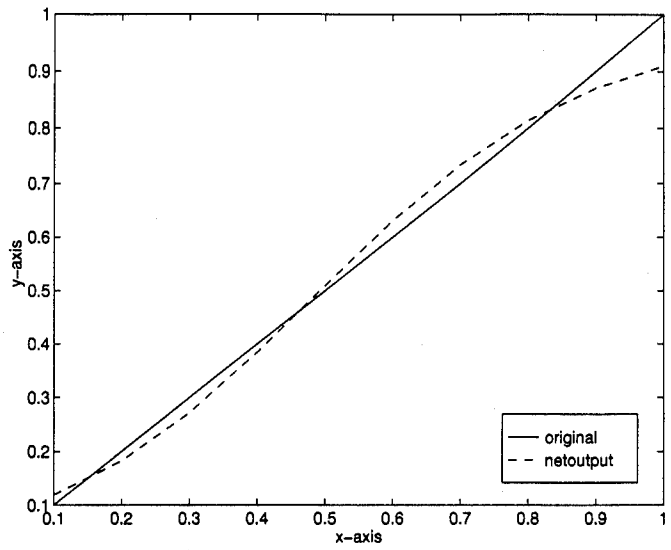


Figure 21. $y = x$ (solid) and MLP (1:10:1) output (dash) after 1000 epochs

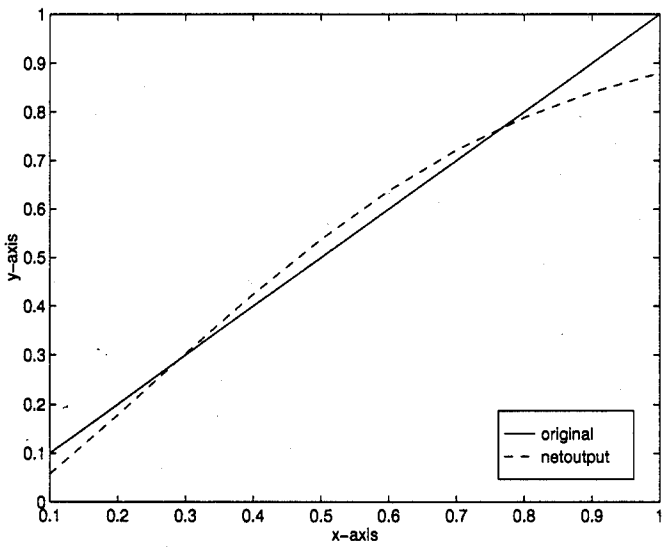


Figure 22. $y = x$ (solid) and MLP (1:10:1) output (dash) after 300 epochs

y	x
.1	.1
.4	.2
.6	.3
.9	.4
1.0	.5
1.0	.6
.8	.7
.5	.8
.3	.9
.1	1.0

Table 2. Simple nonlinear function

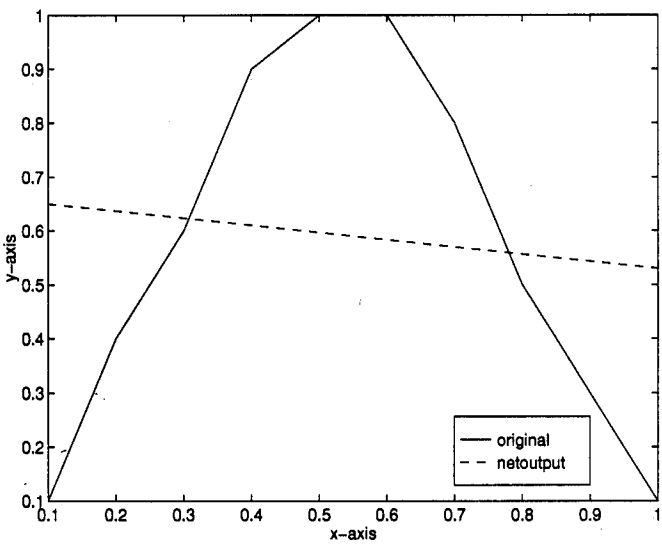


Figure 23. $y \approx -x^2$ (solid) and MLP (1:10:1) output (dash)

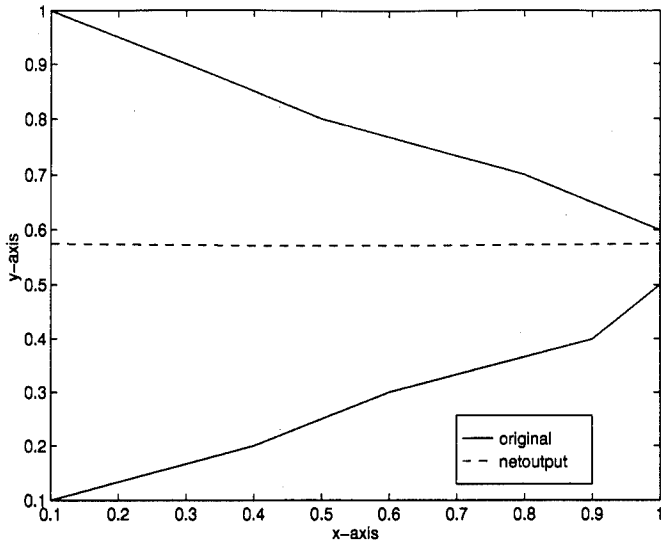


Figure 24. $x \approx -y^2$ (solid) and MLP (1:10:1) output (dash)

data presented in a ASCII matrix where the first columns were the input parameters and the last columns were the output parameters. This is contrary to LNKnet where the first columns are the class. A correction to the columns was made and training of a number of net configuration was resumed. The net could only match a straight line to the data which contributed to the low number of data points.

2.8 Learning LNKmap with a Cosine function

After determining the correct file format, a trigonometric function was tested. The function to be trained was $y = \cos(2\pi x)$, where x ranged from 0 to 1 with a step size of $\frac{1}{629}$ or $\frac{1}{2\pi 100}$ for a total of 629 points. Using the same methodology as before, starting with a net architecture of 1:1:1 and linear output node function many configurations were tried settling on two hidden layers with 10 nodes each (1:10:10:1). Through all these trials, the current experiment was always continued and randomly presented the data. Although, as mentioned above, the data should be presented in order, the strategy at this point in the investigation of LNKmap was to follow the same procedures as LNKnet. The results of the training was that the net learned the first half of the cosine curve and then proceeded asymptotically to a

line parallel to the x axis for the second half of the curve. The approximate number of epochs at this point was 10,000 to 20,000.

At this point it was concluded that random presentation of the data was not working and the data was presented in the order of the input file (x increasing from zero to one). The net learned the function.

In order to duplicate the results, Capt Smith used the same cosine training file. He trained the net for 3000 epochs using random presentation of the data. The net learned the first half as above. Next he continued to train the net to present the data in order. The net learned the whole cosine function. Unfortunately, plots of this data are not available because the LNKmap program has since been changed as discussed below.

2.8.1 Cosine Revisited. From early training attempts with LNKmap, it was determined that LNKmap did not correctly train on data with more than one input. The author of LNKmap was contacted and a fix was obtained. After this fix, the cosine function was retrained using both a RBF and a MLP architecture. This time the function trained was $y = \cos(x)$, where x ranged for 0 to 2π with a step size of 0.01 for a total of 629 points (see figure 25). Before training, the data was normalized to values between 0 and 1 using a Dynamic Range Compression (DRC) algorithm (see figure 26). Given data in a matrix where each column represents a variable, the DRC subtracts the minimum value for each column from each value in the column and divides by the maximum value of the difference for each column. The Matlab code can be seen in appendix C.

2.8.1.1 RBF Architecture. The first architecture used for training was a Radial Basis Function (RBF) using a K-means clustering algorithm. The RBF was trained using 1, 2, 8, 9 and 32 centers in the K-means clustering. Figure 27 shows the result with 1 center in the K-means clustering. Figure 28 shows the result with 2 centers in the K-means clustering. Since the cosine function in the range 0 to 2π

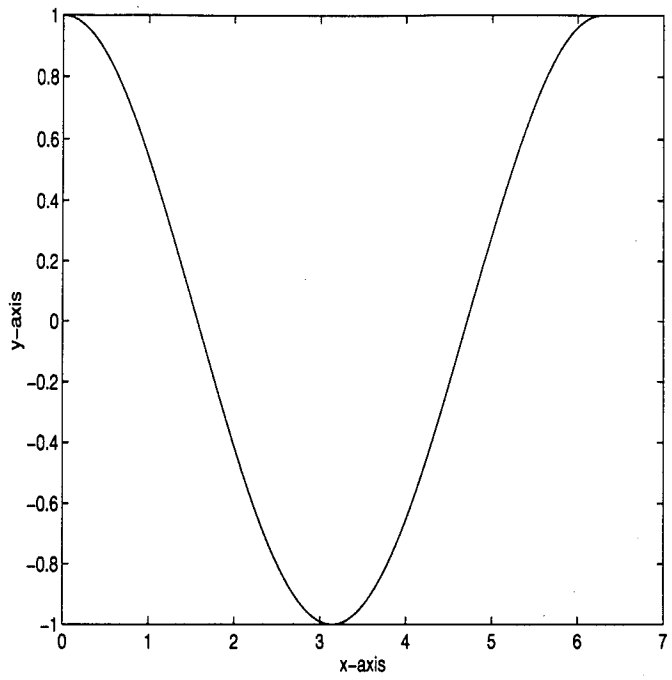


Figure 25. Training function $y = \cos(x)$, where x ranged for 0 to 2π with a step size of 0.01 for a total of 629 points

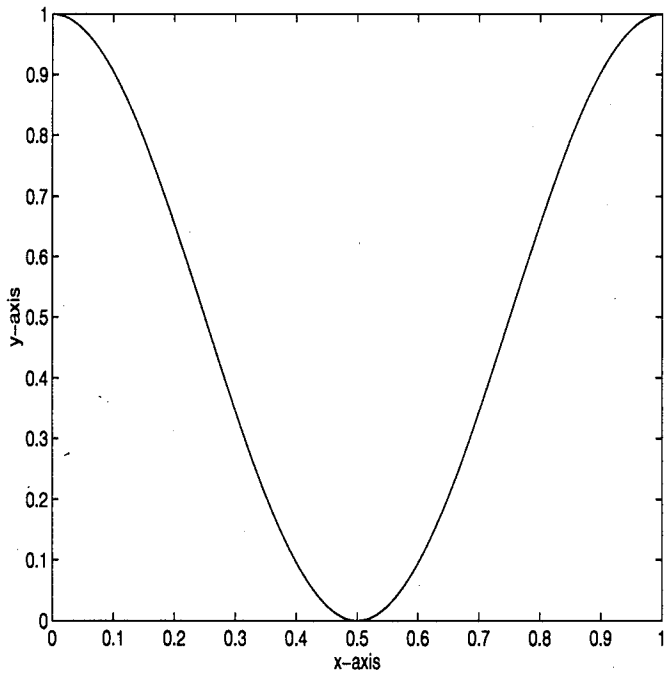


Figure 26. Dynamic Range Compression of $y = \cos(x)$ to values between 0 and 1

is essentially unimodal only one clustering center is needed. Forcing more than one center increases the error. The LNKmap clustering algorithm does not place centers in the same place. A different clustering algorithm may have placed the centers in the same place. Figure 29 shows that adding more centers will reduce the error, but at the expense of a larger network.

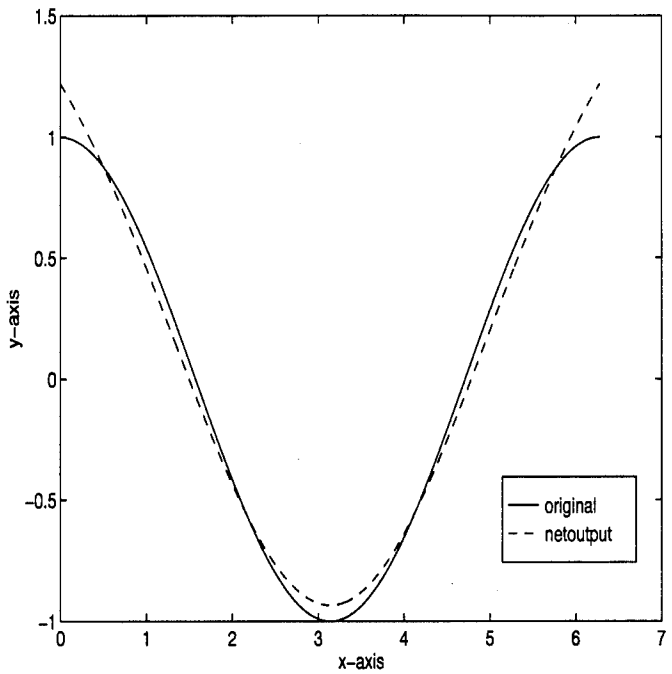


Figure 27. $y = \cos(x)$ (solid) and RBF (1:1:1) output (dash) with 1 center in the K-means clustering algorithm

2.8.1.2 MLP Architecture. Next a MLP architecture of 1:10:10:1 was chosen to match the best results from the previous cosine training (before the software fix). The MLP learned the cosine in only 220 epochs. Figure 30 shows the results after training for 15, 100 and 220 epochs respectively.

2.8.2 Conclusion. It has not been determined why the network needed the data presented first randomly and then in order to learn the function. Presenting the data only randomly or only in order did not produce satisfactory results. The obvious answer is that a bug was in the program. After the fix the program performed as

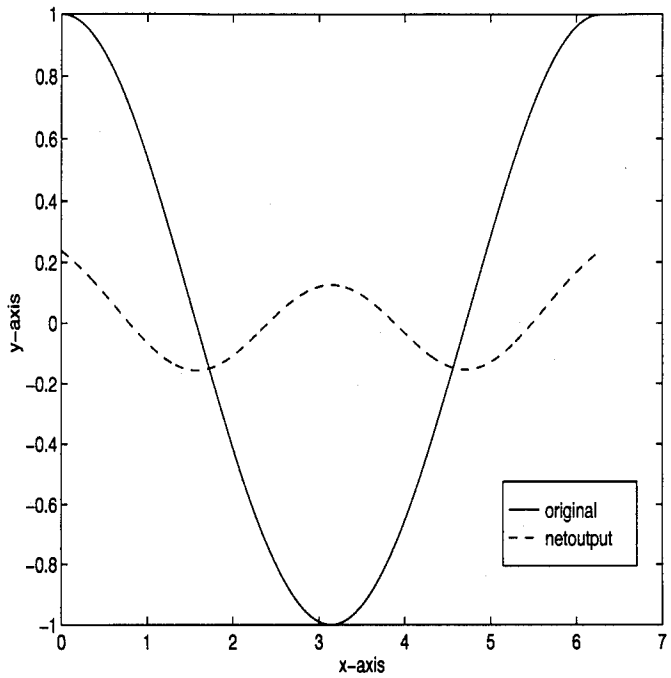


Figure 28. $y = \cos(x)$ (solid) and RBF (1:2:1) output (dash) with 2 centers in the K-means clustering algorithm

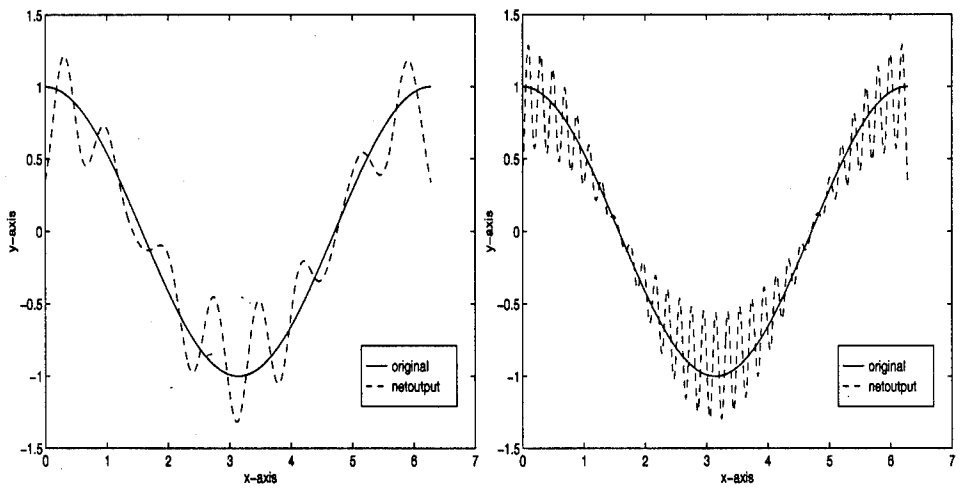


Figure 29. $y = \cos(x)$ (solid) and RBF (1:9:1, 1:32:1) output (dash) with 9 and 32 centers in the K-means clustering algorithm

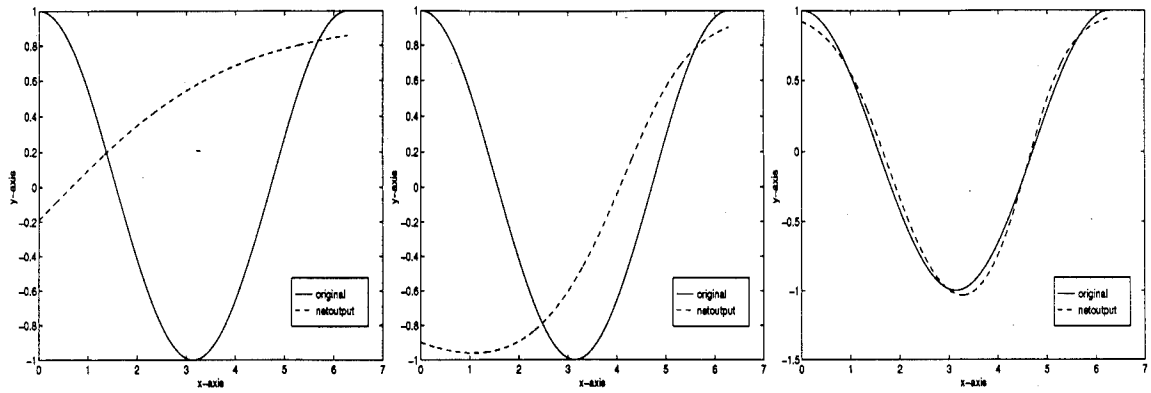


Figure 30. $y = \cos(x)$ (solid) and MLP (1:10:10:1) output (dash) after 15, 100 and 220 epochs

expected. Meaning when the data was presented in order the network learned the function faster. Regardless, the major effort of this investigation was to learn how to use LNKmap as a function approximation tool. This task was accomplished.

2.9 Learning LNKmap with a Two Input Cosine function

The next function tested was the "bird function," $z = \cos(2\pi(x^2 - y^2))$ where x and y ranged from 0 to 1 for a total of 10201 training points (see figure 31). This

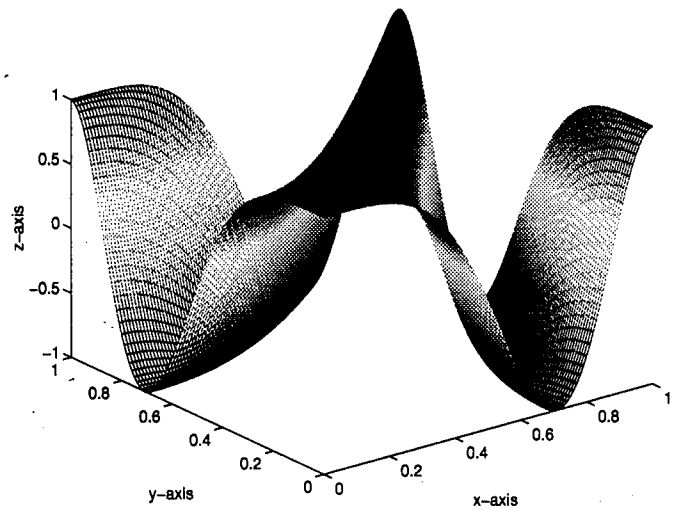


Figure 31. $z = \cos(2\pi(x^2 - y^2))$ where x and y ranged from 0 to 1.

function was chosen because of its simplicity and the output values were relatively

low between -1 and 1 so no normalization was required. Expanding the axes of the bird function actually shows the whole function to be a symmetric cross shape (see figure 32).

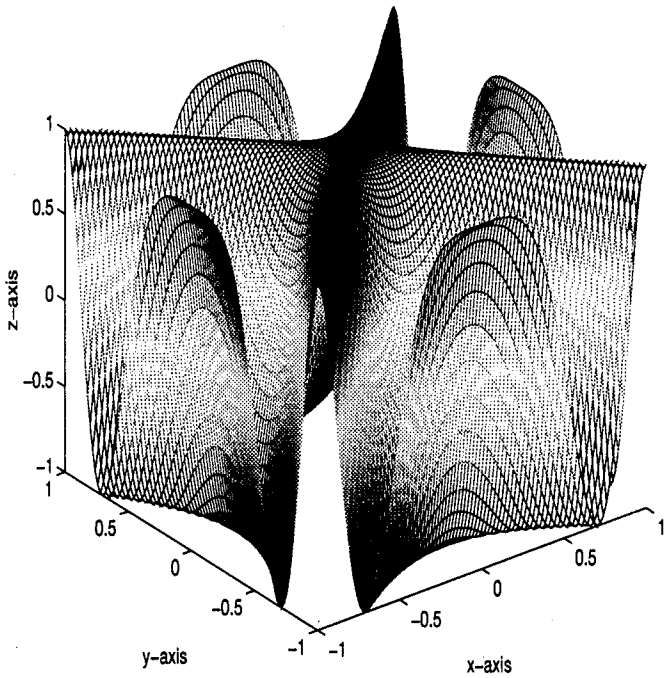


Figure 32. $z = \cos(2\pi(x^2 - y^2))$ where x and y ranged from -1 to 1.

2.9.1 RBF Architecture. Two network architectures were tested with the bird data. The first was a Radial Basis Function (RBF) using a K-means clustering algorithm. The RBF was trained using 8, 64 and 128 centers in the K-means clustering (see figure 33). Power of two centers above 128 (i.e. 256, 512, etc...) caused an error in the program. The best the LNKmap RBF algorithm could learn the bird data is shown in figure 33. The placement of the clustering centers can be seen as bumps in the bird's back. The RMS error for the training using 8, 64 and 128 centers is shown in figure 34. The highest error was seen near the edge of the graphs where the function was constrained. Increasing the number of centers in this case did not reduce the error. The mean error for the 64 centers case was 0.1229 while the mean error for the 128 centers case was 0.1273. Figures 33 and 34 do not show

much difference between the 64 center and 128 center cases. To examine the effect of doubling the number of clustering centers from 64 to 128, the difference between the two outputs is shown in figure 35. The average difference was 9.3888×10^{-5}. The difference between the two RMS errors is shown in figure 35. The average difference of the RMS error was 0.0044.

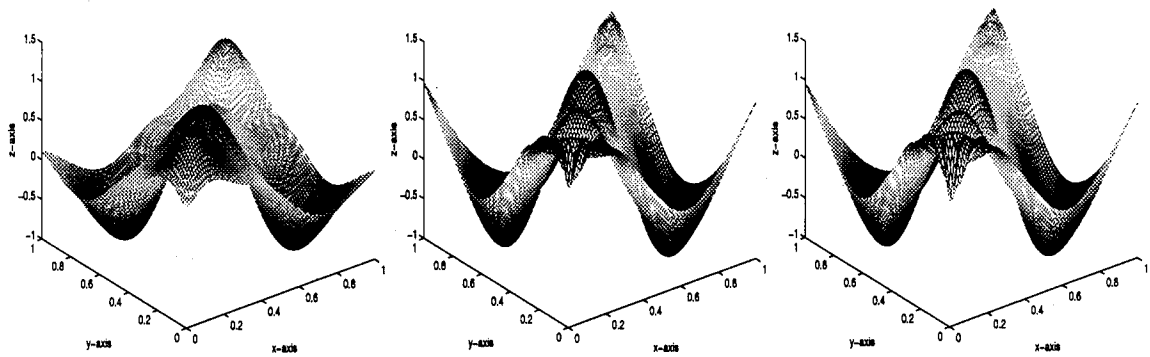


Figure 33. Output of RBF with 8, 64 and 128 centers in the K-means clustering algorithm

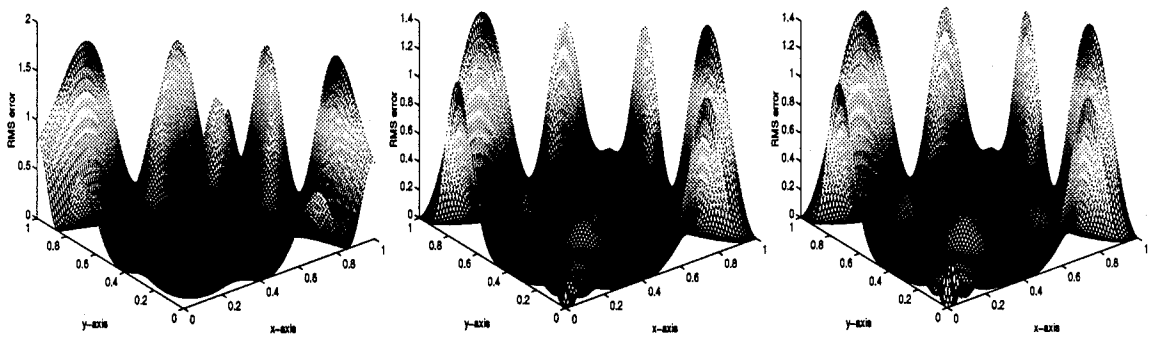


Figure 34. RMS error of RBF with 8, 64 and 128 centers in the K-means clustering algorithm

2.9.2 MLP Architecture. The second network architecture used to train on the bird function was a Multi-Layer Perceptron (MLP). Three MLP architectures were chosen to train the bird data. For each architecture, the non-linearity function used was the symmetric sigmoid (see figure 36). The first architecture had one hidden

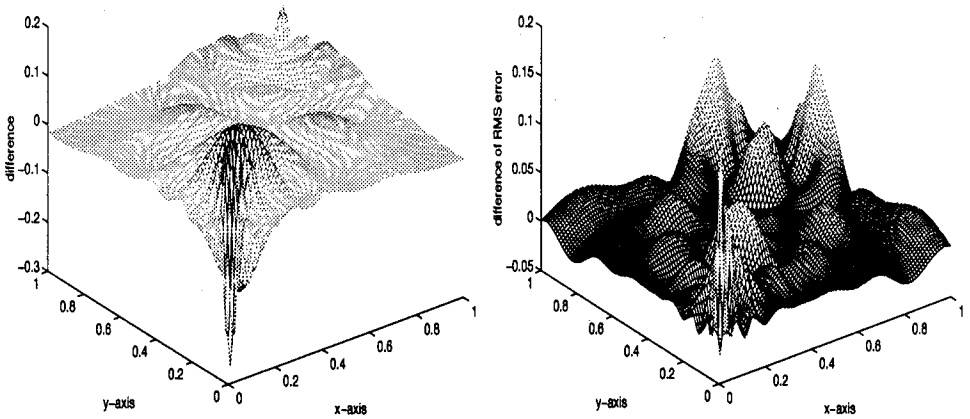


Figure 35. Difference between outputs (left) and RMS errors (right) using 64 and 128 centers in the K-means clustering algorithm

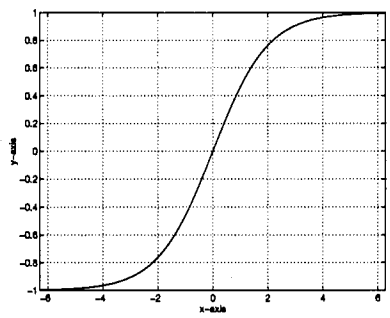


Figure 36. Symmetric sigmoid used for non-linearity function $y = \frac{2}{1+e^x} - 1$

layer of 10 nodes (2:10:1). Figure 37 shows the progress of learning for 220 and 440 epochs respectively. The RMS error is shown in figure 38. The second architecture

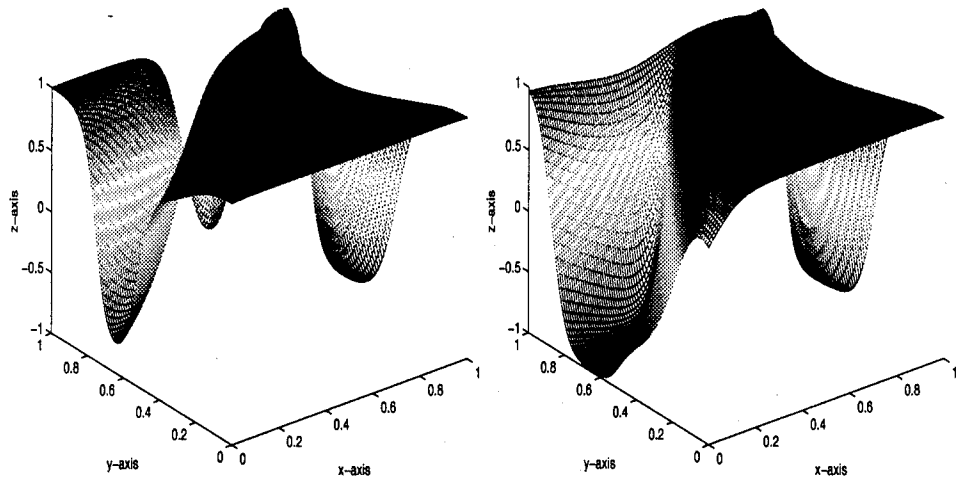


Figure 37. Output of MLP after 220 and 440 epochs with architecture of 2:10:1

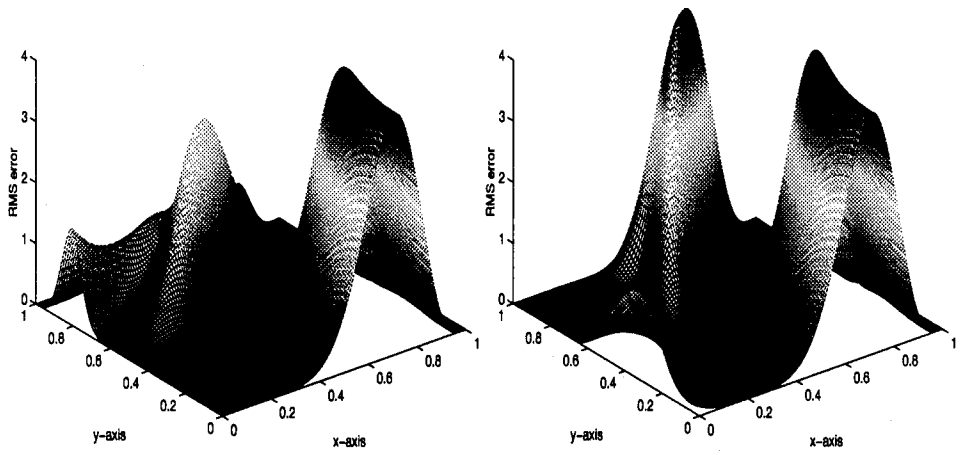


Figure 38. RMS error of MLP after 220 and 440 epochs with architecture of 2:10:1

had one hidden layer of 64 nodes (2:64:1). This architecture was chosen to compare with the 64 center RBF. Figure 39 shows the progress of learning for 20, 50 and 100 epochs respectively. The RMS error is shown in figure 40. Comparing figures 33 and 39 it is clear that the RBF outperformed this architecture in both speed and error even though the architectures have basically the same number of nodes. The third

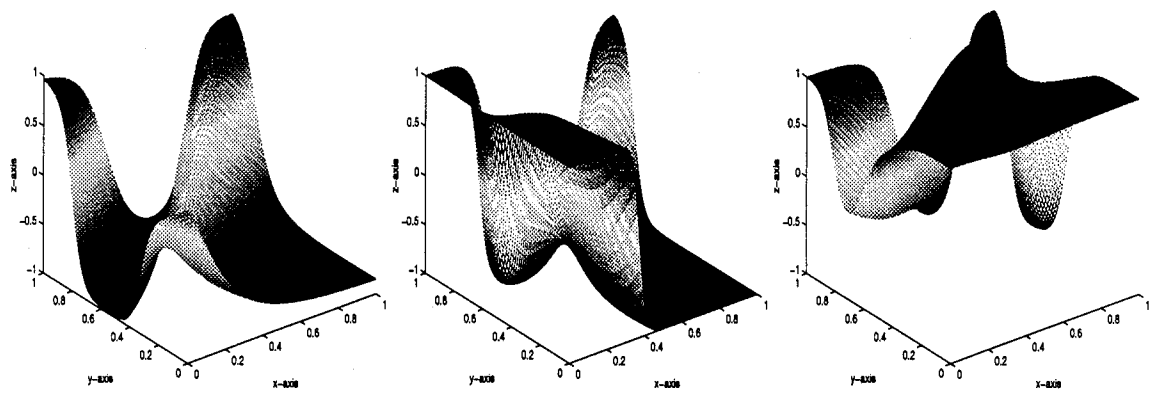


Figure 39. Output of MLP after 20, 50 and 100 epochs with architecture of 2:64:1

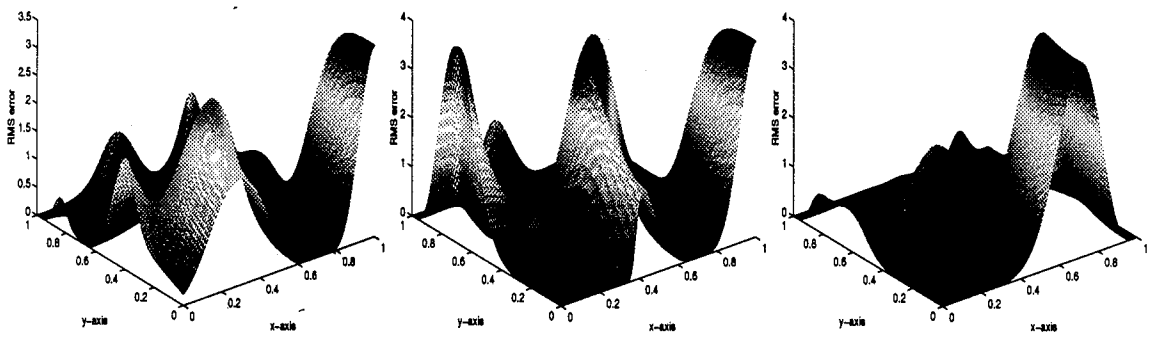


Figure 40. RMS error of MLP after 20, 50 and 100 epochs with architecture of 2:64:1

architecture had two hidden layers of 10 nodes each (2:10:10:1). Figure 41 shows the progress of learning for 20, 220 and 440 epochs respectively. The RMS error is shown in figure 42.

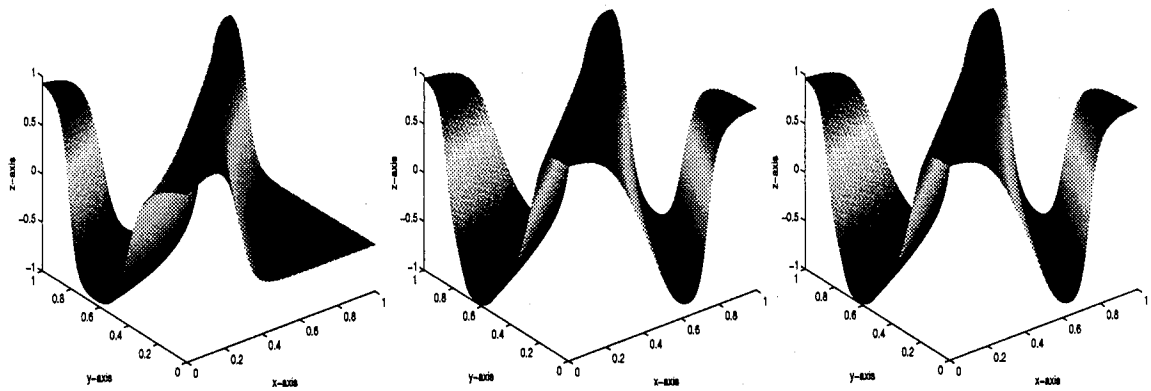


Figure 41. Output of MLP after 20, 220 and 440 epochs with architecture of 2:10:10:1

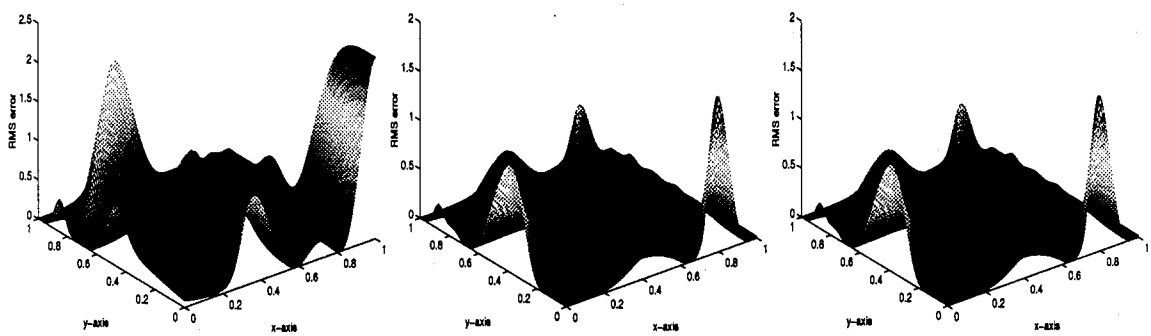


Figure 42. RMS error of MLP after 20, 220 and 440 epochs with architecture of 2:10:10:1

2.9.3 Conclusion. Comparing figures 33 to 41 with figure 31 (the original bird data), it is clear from these results that the MLP was able to learn the bird data better than the RBF. However it is important to note that this result does not generalize to other data and the different architectures produced alternate results. The main purpose of the bird data tests was to verify that the LNKmap program could approximate a function with more than one input.

2.10 Conclusion

This chapter explored the background of both actual and simulated 3D sound. It showed that natural errors occur (i.e. reversals) with both actual 3D and simulated 3D sound sources. The head, pinna and HRTF are important for 3D sound localization. It also showed that non-individualized HRTFs can be successfully used with speech.

The chapter also explored the use of LNKmap as a tool for function approximation. LNKmap was able to successfully train a number of different neural network architectures from simple linear and non-linear functions to two input trigonometric functions.

III. Methodology

3.1 Introduction

The previous chapters explored the background of 3D sound and neural networks. This chapter will combine the two concepts to train a neural network to approximate the 3D sound HRTF.

3.2 Learning HRTF Data for Zero Elevation

3.2.1 MLP Architecture. Having shown that the LNKmap program works correctly for multiple input data, the next step is to map the Head Related Transfer Function (HRTF) data. The HRTF data consists of 25,296 data points. Broken down that is 272 azimuth and elevation locations: 10^o to 350^o in azimuth and -90^o to 90^o in elevation. For each azimuth and elevation pair, 93 frequencies ranging from 100 to 20,000 Hz are used. To simplify the training, only the zero elevation is used in this thesis. The zero elevation data consists of 24 azimuth points with 93 frequencies each. For a total of 2232 data points (see figure 43). Since the MLP worked the

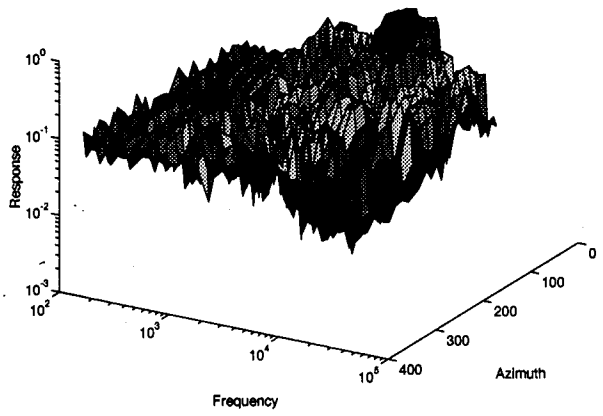


Figure 43. Zero elevation HRTF data for the left ear. Frequency and Response shown in logarithm scale

best with the bird data, the zero elevation was first trained with a MLP with an architecture of 2:50:50:1. The output node function was a symmetric sigmoid. The weights were updated after each trial (no batch). Figure 44 shows the results after 7235 epochs. The MLP extremely smoothes the HRTF data. An experiment needs

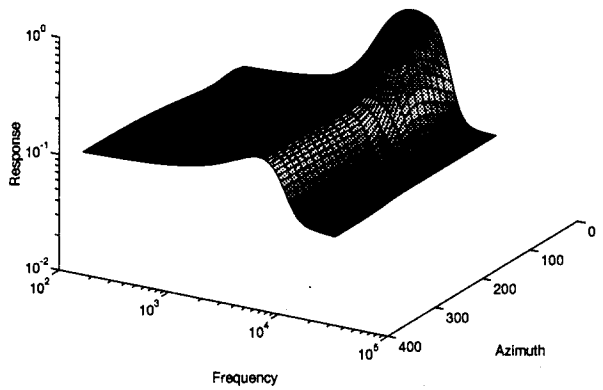


Figure 44. Zero elevation MLP (2:50:50:1) after 7235 epochs. Frequency and Response shown in logarithm scale

to be constructed to determine if the jaggedness of the actual HRTF data is needed to perceive 3D sound. This is not addressed in this thesis.

3.2.2 RBF Architecture. The RBF was tested with the zero elevation data next. A K-means clustering algorithm with 256 centers was used. Compared with the MLP the RBF trained a lot faster and results were even better from a visual standpoint. Figure 45 shows the results of training with the RBF. Blank areas in the figure are where the RBF output was negative or zero (logarithmic scale requires positive values). Since negative HRTFs are not possible, they will be zeroed for filter design. The RBF is able to more closely match the jaggedness of the actual HRTF data. Figure 46 show the RMS error for the MLP and RBF networks respectively. A close inspection of figure 46 reveals that the maximum magnitude for RBF error (0.2707) was actually higher than the MLP error (0.1573). However the MLP had

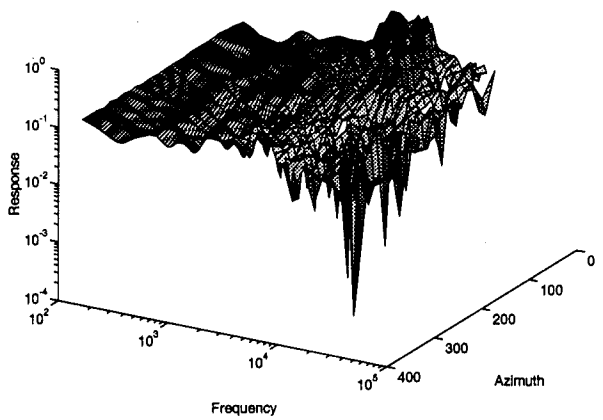


Figure 45. Zero elevation RBF with 256 centers. Frequency and Response shown in logarithm scale

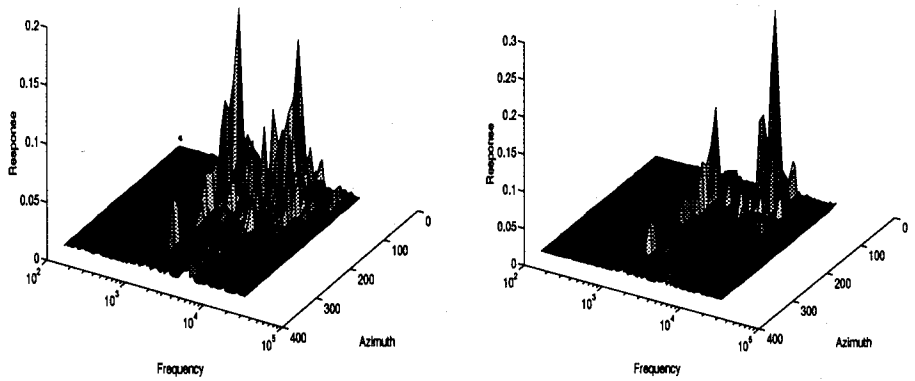


Figure 46. Zero elevation RMS err for the MLP (left) and RBF (right)

more error overall. The sum of the MLP and RBF error was 7.3518 and 5.8528 respectively.

3.2.3 Conclusion. It would be premature to generalize that the RBF works better than the MLP for jagged data like the HRTF data as compared to smooth data like the bird data. The MLP and RBF had two different architectures. The number of weights used by the MLP used was 100 weights (50 weights per hidden node) while the RBF used 512 weights (256 centers by K-means and 256 weights). That is a ratio of 5 to 1. However the RBF does have the desirable feature of speed.

3.3 Experiment 1: With and Without HRTF

In the process of manipulating and filtering sound sources with the HRTF an obvious question arose, "Whether or not the HRTF data helped in 3D localization?" To answer this question experiment one was conceived.

3.3.1 Objective. Determine whether or not the utilization of AAMRL's HRTF data for filter design provides a specific advantage over the use of only ITD's in generation of a binaural signal. The filters were designed using the program ESPS by Entropic Research lab. The design was an FIR filter using the weighted mean square error criterion. The weights corresponded to the AAMRL HRTF for the 93 frequencies at the test location. A separate filter was designed for each ear and for each location tested.

3.3.2 Method. Twenty subjects were presented 48 binaural sounds through a set of headphones. In this forced choice experiment, the subjects were asked to select the location from which they perceived the sound. Subjects choose from a display provided on a computer screen by moving the mouse to place the cursor over the perceived location and clicking the button (see figure 47). The display had only 24 possible azimuths to choose from, and the elevation angle is always zero. These

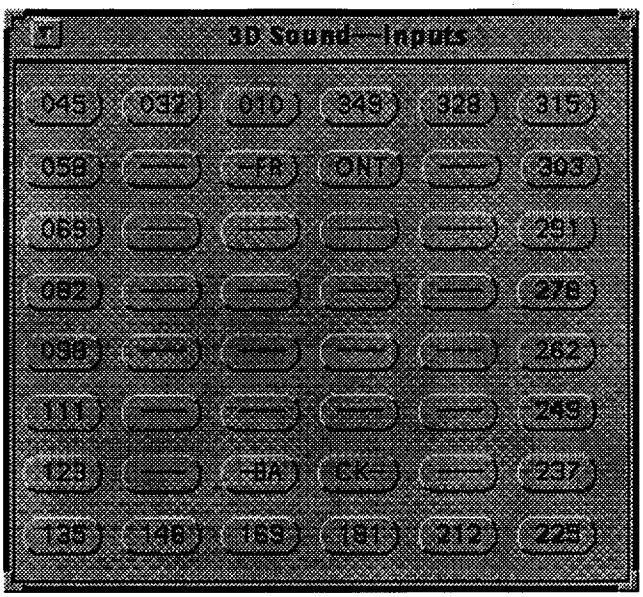


Figure 47. Display provided on a computer screen for selection of the perceived location

24 locations corresponded to the 24 possible locations of the computer generated binaural signals. The signal used in each case was either digitally filtered and delayed appropriately for a given angle of azimuth, or simply delayed appropriately to account for the angle. In either case, the original signal was the speech pattern for a male utterance of the word "seventeen."

Subjects were asked to close their eyes during the data collection phase of this experiment. Interaction effects between the subjects and the two fixed factors (angle and method of signal generation) in this experiment are assumed to be zero. The angle at which the signal is designed to come from was called factor B, and has 24 levels (one for each azimuth). Factor A represented the method by which the binaural signal was generated, and has 2 levels (one for HRTF combined with ITD and one for ITD only). Factor C represented the subjects.

3.3.3 The Model. The model used in experiment one is an analysis of variance (ANOVA) model for a reduced three-factor experiment (14). It is assumed that the $a \cdot b \cdot c$ treatment combinations represent random samples of size n, where

$a = 2$, the number of filters; $b = 24$, the number of azimuths; $c = 20$, the number of subjects; and $n = 1$, the sample size. The model is

$$X_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \varepsilon_{ijkl} \tag{1}$$

where

$$i = 1, 2, \ldots, a$$

$$j = 1, 2, \ldots, b$$

$$k = 1, 2, \ldots, c$$

$$l = 1, 2, \ldots, n$$

The X_{ijkl} are assumed to be normally distributed and the following restrictions apply:

$$\sum_i \alpha_i = 0, \quad \sum_j \beta_j = 0, \quad \sum_k \gamma_k = 0 \tag{2}$$

$$\sum_i (\alpha\beta)_{ij} = 0, \quad \sum_j (\alpha\beta)_{ij} = 0 \tag{3}$$

$$\sum_i (\alpha\gamma)_{ik} = 0, \quad \sum_k (\alpha\gamma)_{ik} = 0 \tag{4}$$

$$\sum_j (\beta\gamma)_{jk} = 0, \quad \sum_k (\beta\gamma)_{jk} = 0 \tag{5}$$

Here α_i, β_j, and γ_k represent the main effects due to factors filter, azimuth and subject respectively. The terms $(\alpha\beta)_{ij}$, $(\alpha\gamma)_{ik}$, and $(\beta\gamma)_{jk}$ represent interaction between the factors. Let

$\mu_{ijk\cdot}$ = mean for the (ijk)th treatment combination

$\mu_{ij\cdot\cdot}$ = the average of the population means for those populations receiving the ith level of A and

the jth level of B

$$= \sum_{k=1}^{c} \frac{\mu_{ijk\cdot}}{c} \tag{6}$$

$\mu_{i\cdots}$ = the average of the population means for those population receiving the ith level of A

$$= \sum_{j=1}^{b}\sum_{k=1}^{c} \frac{\mu_{ijk\cdot}}{bc} \tag{7}$$

μ = the average of the population means for all populations under consideration

$$= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} \frac{\mu_{ijk\cdot}}{abc} \tag{8}$$

(9)

The other mean terms are defined in a similar manner:

$$\mu_{i\cdot k\cdot} = \sum_{j=1}^{b} \frac{\mu_{ijk\cdot}}{b} \tag{10}$$

$$\mu_{\cdot jk\cdot} = \sum_{i=1}^{a} \frac{\mu_{ijk\cdot}}{a} \tag{11}$$

$$\mu_{\cdot j\cdot\cdot} = \sum_{i=1}^{a}\sum_{k=1}^{c} \frac{\mu_{ijk\cdot}}{ac} \tag{12}$$

$$\mu_{\cdot\cdot k\cdot} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{\mu_{ijk\cdot}}{ab} \tag{13}$$

(14)

The main effects and interaction effects are defined as follows:

$$\alpha_i = \mu_{i\cdots} - \mu \tag{15}$$

$$\beta_j = \mu_{\cdot j\cdot\cdot} - \mu \tag{16}$$

$$\gamma_k = \mu_{\cdot\cdot k\cdot} - \mu \tag{17}$$

$$(\alpha\beta)_{ij} = \mu_{ij\cdot\cdot} - \mu_{i\cdots} - \mu_{\cdot j\cdot\cdot} + \mu \tag{18}$$

$$(\alpha\gamma)_{ik} = \mu_{i\cdot k\cdot} - \mu_{i\cdots} - \mu_{\cdot\cdot k\cdot} + \mu \tag{19}$$

$$(\beta\gamma)_{jk} = \mu_{jk\cdot\cdot} - \mu_{\cdot j\cdot\cdot} - \mu_{\cdot\cdot k\cdot} + \mu \tag{20}$$

The totals for the three-way ANOVA are shown next:

$T_{i\cdots}$ = sum of all responses from experimental units receiving level i of A and B; where x is the absolute difference of the actual location and perceived location of the sound

$$= \sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n} x_{ijkl} \tag{21}$$

$T_{ij\cdot\cdot}$ = sum of all responses from experiment units receiving level i of A and j of B

$$= \sum_{k=1}^{c}\sum_{l=1}^{n} x_{ijkl} \tag{22}$$

The other total terms are defined in a similar manner:

$$T_{\cdot j\cdot\cdot} = \sum_{i=1}^{a}\sum_{k=1}^{c}\sum_{l=1}^{n} x_{ijkl} \tag{23}$$

$$T_{\cdot\cdot k\cdot} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{l=1}^{n} x_{ijkl} \tag{24}$$

$$T_{i\cdot k\cdot} = \sum_{j=1}^{b}\sum_{l=1}^{n} x_{ijkl} \tag{25}$$

$$T_{\cdot jk\cdot} = \sum_{i=1}^{a}\sum_{l=1}^{n} x_{ijkl} \tag{26}$$

$$T_{ijk\cdot} = \sum_{l=1}^{n} x_{ijkl} \tag{27}$$

$$T_{\cdots\cdot} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n} x_{ijkl} \tag{28}$$

The sums of squares needed in the three-way ANOVA are:

$$SS_A = \sum_{i=1}^{a} \frac{T_{i\cdots}^2}{bcn} - \frac{T_{\cdots\cdot}^2}{abcn} \tag{29}$$

$$SS_B = \sum_{j=1}^{b} \frac{T_{\cdot j \cdot \cdot}^2}{acn} - \frac{T_{\cdots\cdot}^2}{abcn} \tag{30}$$

$$SS_C = \sum_{k=1}^{c} \frac{T_{\cdot\cdot k \cdot}^2}{abn} - \frac{T_{\cdots\cdot}^2}{abcn} \tag{31}$$

$$SS_{AB} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{T_{ij\cdot\cdot}^2}{cn} - \sum_{i=1}^{a} \frac{T_{i\cdots}^2}{bcn} - \sum_{j=1}^{b} \frac{T_{\cdot j\cdot\cdot}^2}{acn} + \frac{T_{\cdots\cdot}^2}{abcn} \tag{32}$$

$$SS_{AC} = \sum_{i=1}^{a}\sum_{k=1}^{c} \frac{T_{i\cdot k\cdot}^2}{bn} - \sum_{i=1}^{a} \frac{T_{i\cdots}^2}{bcn} - \sum_{k=1}^{c} \frac{T_{\cdot\cdot k\cdot}^2}{abn} + \frac{T_{\cdots\cdot}^2}{abcn} \tag{33}$$

$$SS_{BC} = \sum_{j=1}^{b}\sum_{k=1}^{c} \frac{T_{\cdot jk\cdot}^2}{an} - \sum_{j=1}^{b} \frac{T_{\cdot j\cdot\cdot}^2}{acn} - \sum_{k=1}^{c} \frac{T_{\cdot\cdot k\cdot}^2}{abn} + \frac{T_{\cdots\cdot}^2}{abcn} \tag{34}$$

$$SS_{\text{total}} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n} x_{ijkl}^2 - \frac{T_{\cdots\cdot}^2}{abcn} \tag{35}$$

$$SS_E = SS_{\text{total}} - SS_A - SS_B - SS_C - SS_{AB} - SS_{AC} - SS_{BC} \tag{36}$$

Finally an ANOVA table as shown in table 3 is set up to test the null (H_0) and alternative (H_1) hypothesis:

$H_0^{(1)}$: There is no interactions of the factors A and B means; that is, $(\alpha\beta)_{ij} = 0$.
$H_0^{(2)}$: There is no interactions of the factors A and C means; that is, $(\alpha\gamma)_{ik} = 0$.
$H_0^{(3)}$: There is no interactions of the factors B and C means; that is, $(\beta\gamma)_{jk} = 0$.
$H_0^{(4)}$: Factor A means are equal; that is, $\alpha_i = 0$.
$H_0^{(5)}$: Factor B means are equal; that is, $\beta_j = 0$.
$H_0^{(6)}$: Factor C means are equal; that is, $\gamma_k = 0$.
$H_1^{(1-6)}$: The means are not equal.

3.4 Experiment 2: AAMRL HRTF and RBF HRTF

3.4.1 Objective. Determine whether or not the utilization of AAMRL's HRTF data for filter design is statistically different from the utilization of RBF neural network data for filter design in generation of a binaural signal.

Source of Variation	Degrees of Freedom	Sum of Squares (SS)	Mean square (MS)	Calculated F Value
A	$a-1$	SS_A	$SS_A/(a-1)$	MS_A/MS_E
B	$b-1$	SS_B	$SS_B/(b-1)$	MS_B/MS_E
C	$c-1$	SS_C	$SS_C/(c-1)$	MS_C/MS_E
AB	$(a-1)(b-1)$	SS_{AB}	$SS_{AB}/(a-1)(b-1)$	MS_{AB}/MS_E
AC	$(a-1)(c-1)$	SS_{AC}	$SS_{AC}/(a-1)(c-1)$	MS_{AC}/MS_E
BC	$(b-1)(c-1)$	SS_{BC}	$SS_{BC}/(b-1)(c-1)$	MS_{BC}/MS_E
Error (E)	DF_E = subtraction	SS_E	SS_E/DF_E	
Total	$abcn-1$	SS_{total}		

Table 3. Analysis of Variance for "Reduced three-way model"

3.4.2 Method. Twenty subjects were presented 48 binaural sounds through a set of headphones. In this forced choice experiment, the subjects were asked to select the location from which they perceived the sound to come from. Subjects chose from a display provided on a computer screen by moving the mouse to place the cursor over the perceived location and clicking the button (see figure 47).

The display had only 24 possible azimuths to choose from, and the elevation angle is always zero. These 24 locations corresponded to the 24 possible locations of the computer generated binaural signals. The signal used in each case are either digitally filtered with the AAMRL HRTF data or RBF HRTF data and delayed appropriately for a given angle of azimuth. The filters were designed using the program ESPS by Entropic Research lab. The design was an FIR filter using the weighted mean square error criterion. The weights corresponded to the AAMRL HRTF or RBF HRTF for the 93 frequencies at the test location. A separate filter was designed for each ear and for each location tested. In either case, the original signal is the speech pattern for a male utterance of the word "seventeen."

Subjects were asked to close their eyes during the data collection phase of this experiment. Interaction effects between the subjects and the two fixed factors (angle and method of signal generation) in this experiment are assumed to be zero. The angle at which the signals are designed to come from are called factor B, and has

24 levels (one for each azimuth). Factor A represented the method by which the binaural signal was generated, and has 2 levels (one for AAMRL HRTF and one for RBF HRTF). Factor C represented the subjects.

3.4.3 The Model. The model used in experiment two is the same as described above in experiment one.

3.5 Conclusion

This chapter presented the methodology for training a neural network to approximate the HRTF for zero elevation. It also presented two experiments to verify the methodology. The results of the training and experiments will be presented in the following chapter.

IV. Experiment Results

4.1 Introduction

This chapter presents the results of the two experiments conducted during this thesis work. Before presenting the results of each experiment a discussion of the types of errors analyzed in the experiments is presented.

4.2 Types of Error

4.2.1 Algebraic Error. Algebraic error is measured by taking the difference between the actual location of the sound source and the subjects perceived location. The difference is calculated so that the results is always between or equal to $\pm 180^o$. Also the sign of the result indicates the direction the subjects response was from the actual location. Positive values indicate a clockwise difference and negative values indicate a counter-clockwise difference (see figure 48). For example, if the actual location is 349^o and the subject responded with 10^o the algebraic error is calculated as $+21^o$. On the other hand, if the actual location is 10^o and the subject responded with 349^o the algebraic error is calculated as -21^o.

4.2.2 Absolute Error. Absolute error is measured by taking the absolute value of the difference between the actual location of the sound source and the subjects perceived location of the sound source. The difference is calculated so that the result is always less than or equal to 180^o. For example, if the actual location is 349^o and the subject responded with 10^o the absolute error is calculated as 21^o as opposed to 339^o.

4.2.3 Reversals. Reversals are a phenomenon where the sound is perceived flipped front to back or back to front from the actual location. Note this is not a 180^o shift, but a symmetric flip about the horizontal axis though the ears (see figure 49). Figure 50 shows actual occurrence of this phenomena from experiment

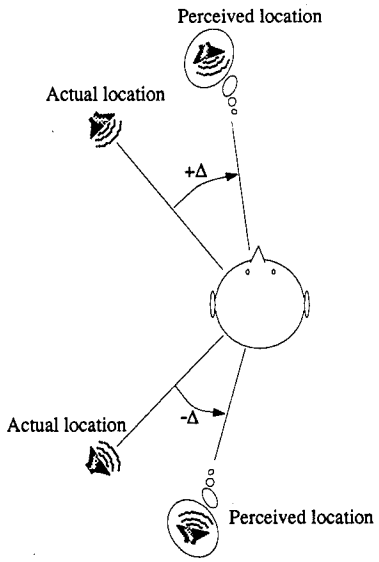


Figure 48. Algebraic Error, positive values indicate a clockwise difference and negative values indicate a counter-clockwise difference

one. The left image shows a subject with back-front reversals with HRTF treatment. The right image shows a different subject with front-back reversals without HRTF treatment. The notation has the following meaning. The × symbol indicates actual location while the ○ symbol indicates perceived location. The line between the × and ○ shows the pairwise groupings. The gradual spacing toward the center of the head is for clarity only and is not an indication of perceived distance.

4.2.4 Default to 90°. Default to 90° is a phenomena where the perceived location defaults to 90° or 270° no matter where the actual location is. This phenomena has been observed by other researchers (17). Figure 51 is an example of a subject whose perceived locations defaulted to 90° and 270°.

4.3 Experiment One Results

As stated previously, the primary objective of experiment one was to determine whether or not the Head Related Transfer Function (HRTF) provides any advantage to a listener attempting to identify the azimuth of a computer generated binaural

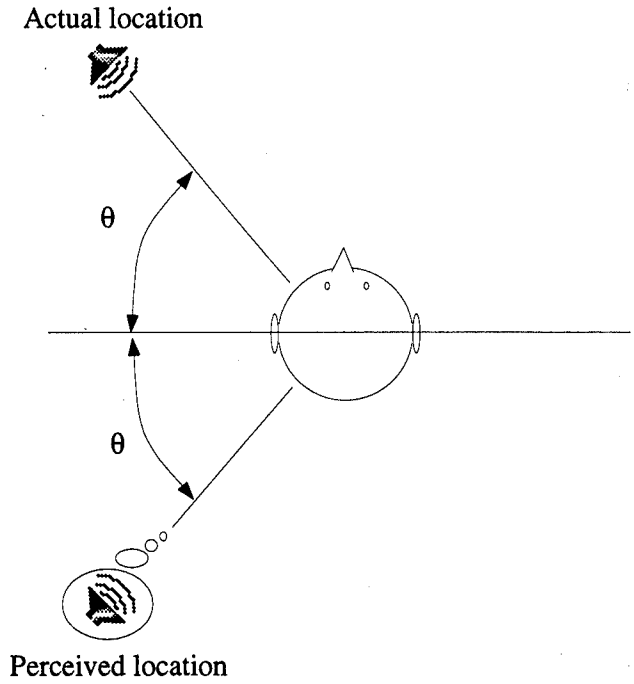


Figure 49. Reversal, a symmetric flip about the horizontal axis through the ears

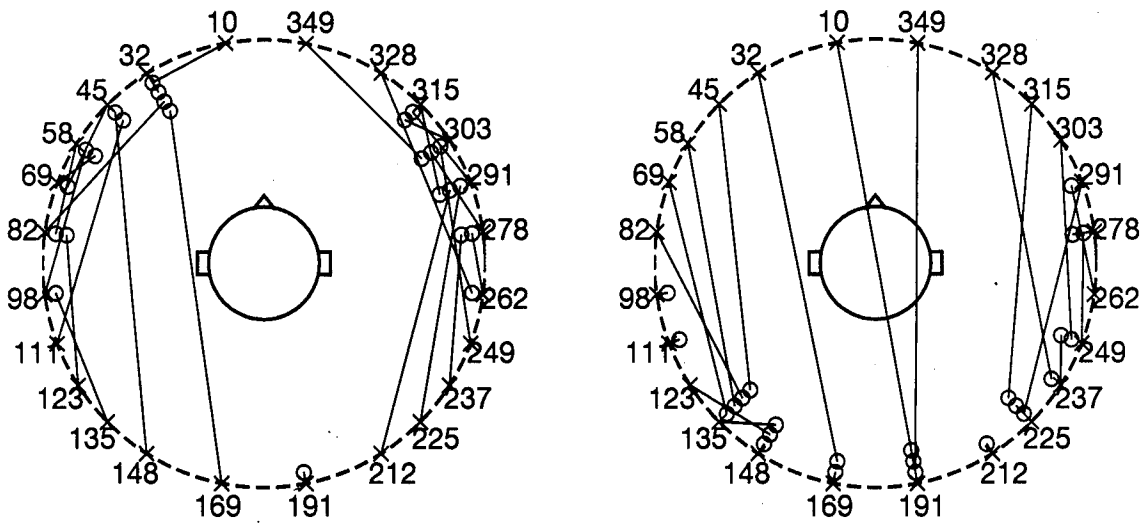


Figure 50. Example of a subject with back-front reversals with HRTF treatment (left) and a different subject with front-back reversals without HRTF treatment (right). Notation: × indicates actual location, ○ indicate perceived location. The line between the × and ○ shows the pairwise groupings. The gradual spacing toward the center of the head is for clarity only and is not an indication of perceived distance.

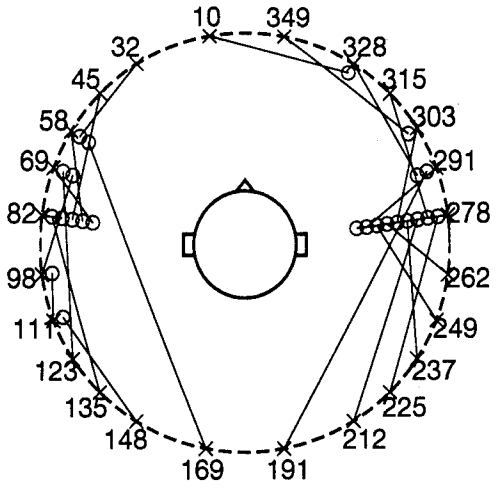


Figure 51. Example of subject whose perceived locations defaulted to 90^o and 270^o. Notation same as above

signal. More specifically, does the use of the HRTF in generating binaural signals provide a listener with a signal that can be more accurately located in azimuth than a binaural signal generated using only interaural time delay (ITD)? The HRTF's used in this experiment were developed at Armstrong Aerospace Medical Research Laboratories (AAMRL), Wright-Patterson AFB, OH (12).

Data collected for 20 subjects was used to determine the answer to the question posed above. Calculations in this statistical analysis of the data are performed using *Matlab*. The data used is found in appendix A. The *Matlab* code is found in appendix C. This experiment is a reduced three factor experiment. Table 4 summarizes the factors in this experiment.

Factor	Levels
Filter	2
Subject	20
Azimuth	24

Table 4. Experiment 1 factors

Where the levels of "Subject" represent each of the twenty individuals taking part in the experiment, and the levels of "Filter" are 1 and 2 (1 corresponds to the

use of the HRTF and 2 corresponds to no use of HRTF). The levels of "Azimuth" are 1, 2, 3, ..., 24 and correspond to the angles shown in table 5. The angle of

Level	Angle	Level	Angle
1	10^o	13	191^o
2	32^o	14	212^o
3	45^o	15	225^o
4	58^o	16	237^o
5	69^o	17	249^o
6	82^o	18	262^o
7	98^o	19	278^o
8	111^o	20	291^o
9	123^o	21	303^o
10	135^o	22	315^o
11	148^o	23	328^o
12	169^o	24	349^o

Table 5. Levels of "Azimuth"

azimuth is measured as shown in figure 52. Using the data set shown in appendix A

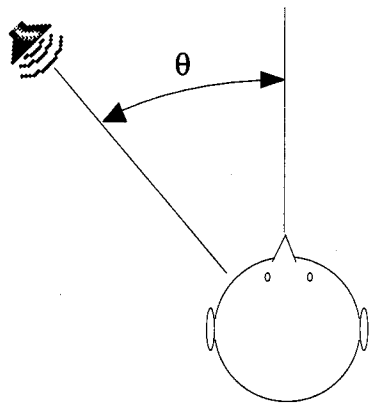


Figure 52. Azimuth, θ of sound source to directly in front of face

with "Response" representing the magnitude of the difference between the subjects' perceived angle and the true location, the following Analysis of Variance (ANOVA) (table 6) was obtained.

Source of Variation	Degrees of Freedom	Sum of Squares (SS)	Mean square (MS)	Calculated F Value	F at 1% level
Filter (A)	1	60554.26	60554.26	57.44 7.47	6.85,6.63 7.88
Azimuth (B)	23	416834.70	18123.25	17.19 2.23	2.03,1.79 2.78,2.70
Subject (C)	19	48451.13	2550.06	2.42 –	2.19,1.88 –
A * B	23	186224.72	8096.73	7.68	2.03,1.79
A * C	19	23134.95	1217.63	1.15	2.19,1.88
B * C	437	564293.18	1291.29	1.22	1.53,1.00
Error	437	460732.48	1054.31		
Total	959	1760225.42			

Table 6. Analysis of Variance for "Response"

Recall, the following model for this experiment:

$$X_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \varepsilon_{ijkl}$$

Interaction effects between filters and azimuth must be tested. The null and alternate hypotheses are $H_0^{(1)}$ (no interactions between filters and azimuth) and H_1 (the means are not equal). The test statistic is F_{calc} = (Mean Square Interaction) $\div$ (Mean Square Random error) = 7.68 (as seen in table 6). Testing at an significance level of 1% ($\alpha = .01$),

$$F_{calc} = 7.68 > F_{.99,23,437} \approx 2.03, 1.79.$$

Hence, $H_0^{(1)}$ can be rejected. Since the hypothesis $H_0^{(1)}$ for the Filter * Azimuth can be rejected, it can be concluded that the interactions are significantly large. Differences in the factors are then significant only if they are large compared with the interactions. For this reason, many statisticians recommend that $H_0^{(4)}$ and $H_0^{(5)}$ be tested by using the F ratios MS_A/MS_{AB} and MS_B/MS_{AB} rather than those given in table 3 (27). These alternate F values are the second values shown in

table 6. Testing $H_0^{(4)}$ at an significance level of 1%,

$$F_{calc_{alt}} = 7.47 \leq F_{.99,1,23} = 7.88.$$

Testing $H_0^{(5)}$ at an significance level of 1%,

$$F_{calc_{alt}} = 2.23 \leq F_{.99,23,23} \approx 2.78, 2.70.$$

Therefore the means of the filters are statistically equal and the means of the azimuths are statistically equal.

Next completing the same tests for Filter $*$ Subject interactions. Testing at an significance level of 1% ($\alpha = .01$),

$$F_{calc} = 1.15 \leq F_{.99,19,437} \approx 2.19, 1.88$$

shows that $H_0^{(2)}$ cannot be rejected. It can be concluded that there is no Filter $*$ Subject interactions and $H_0^{(4)}$ and $H_0^{(6)}$ can be tested. Testing $H_0^{(4)}$ at an significance level of 1%,

$$F_{calc} = 57.44 > F_{.99,1,437} \approx 6.63.$$

Testing $H_0^{(6)}$ at an significance level of 1%,

$$F_{calc} = 2.42 > F_{.99,19,480} \approx 1.91.$$

Therefore the means of the Filters nor the means of the Subjects are statistically equal.

Finally completing the same tests for Azimuth $*$ Subject interactions. Testing at an significance level of 1% ($\alpha = .01$),

$$F_{calc} = 1.22 > F_{.99,437,437} \approx 1.53, 1.00$$

shows that $H_0^{(3)}$ cannot be rejected. Testing $H_0^{(5)}$ at an significance level of 1%,

$$F_{calc} = 17.19 > F_{.99,23,437} \approx 2.03, 1.79.$$

Testing $H_0^{(6)}$ at an significance level of 1%,

$$F_{calc} = 2.42 > F_{.99,19,437} \approx 2.19, 1.88.$$

Therefore the means of the azimuth nor the means of the subjects are statistically equal.

Using the reversal corrections, the following Analysis of Variance (ANOVA) (table 7) was obtained. If the reversal location error was smaller than the subjects' response then the reversal angle was used in table 7. The number of reversals corrected was 432 out of 960 or 45%.

Source of Variation	Degrees of Freedom	Sum of Squares (SS)	Mean square (MS)	Calculated F Value	F at 1% level
Filter (A)	1	1811.7	1811.77	7.45 3.28	6.85,6.63 8.18
Azimuth (B)	23	31188.33	1356.01	5.58 –	2.03,1.79 –
Subject (C)	19	21444.68	1128.67	4.64 2.04	2.19,1.88 3.15,3.00
A * B	23	7804.46	339.32	1.40	2.03,1.79
A * C	19	10498.02	552.53	2.27	2.19,1.88
B * C	437	137136.70	313.81	1.29	1.53,1.00
Error	437	106269.46	243.18		
Total	959	316153.40			

Table 7. Analysis of Variance for "Response" with reversal corrections

From table 7 the calculated F values indicate significant interaction between Filter * Subject. The other interactions Filter * Azimuth and Azimuth * Subject indicate no significant interaction. Comparing within factor variance, the means

of the filters are statistically equal; the means of the azimuths not are statistically equal; and the means of the subjects are statistically equal.

Now, the main objective of this experiment is addressed. Is there a difference in the mean response for HRTF versus non-HRTF binaural signals? To determine the answer to this question, a pairwise comparison of the means for the two levels of the variable "Filter" is conducted. The following formula is used to determine the confidence intervals for the differences of two population means:

$$\bar{X}_2 - \bar{X}_1 \pm z_c \sigma_{\bar{X}_2 - \bar{X}_1} = \bar{X}_2 - \bar{X}_1 \pm z_c \sqrt{\frac{\sigma_2^2}{N_2} + \frac{\sigma_1^2}{N_1}} \tag{37}$$

where,

X_1 is the mean of the with HRTF filter error.
X_2 is the mean of the without HRTF filter error.
σ_1 is the standard deviation of the with HRTF filter error.
σ_2 is the standard deviation of the without HRTF filter error.
z_c is the 99% confidence coefficient ($z_c = 2.58$).
N_1 is the size of the with HRTF filter error.
N_2 is the size of the without HRTF filter error.

And since the point estimates for the means at each level are:

$$\begin{aligned} \bar{X}_1 &= 35.73 \\ \bar{X}_2 &= 51.61 \\ \bar{X}_2 - \bar{X}_1 &= 15.88 \end{aligned}$$

We can state with 99 percent confidence that :

$$8.87 < \bar{X}_2 - \bar{X}_1 < 22.90$$

And since the point estimates for the means at each level with reversal corrections are:

$$\bar{X}_1 = 20.12$$

$$\bar{X}_2 = 17.37$$

$$\bar{X}_2 - \bar{X}_1 = -2.75$$

We can state with 99 percent confidence that :

$$-5.76 < \bar{X}_2 - \bar{X}_1 < 0.27$$

The confidence interval shows a 8.87^o to 22.90^o advantage of using the HRTF versus no HRTF. After reversal correction the advantage is reduced to 0.27^o to a disadvantage of 5.76^o. It should be obvious that the correction for reversals will reduce the error. The algorithm used to include reversals chooses the smaller error. The mean with HRTF error reduced from 35.73^o to 20.12^o and the mean without HRTF error reduced from 51.61^o to 17.37^o. The interaction of reversals is examined further in the next section.

4.3.1 Further examination of results. The above results show mixed results of the statistical advantage in localization accuracy over simply ITD with significant interaction between the location of the sound and whether the HRTF is used.

Figure 53 shows the absolute mean error and standard deviation for sounds with the HRTF (solid) and without the HRTF (dash). The error bars (vertical lines) show the standard deviation of the mean error. This figure reveals that the without HRTF filter had greater error for presentation in front of the head when compared to behind the head. Figure 54 shows the algebraic mean error and standard deviation for sounds with the HRTF (solid) and without the HRTF (dash). Again the error bars show the standard deviation of the mean error. The algebraic error gives a

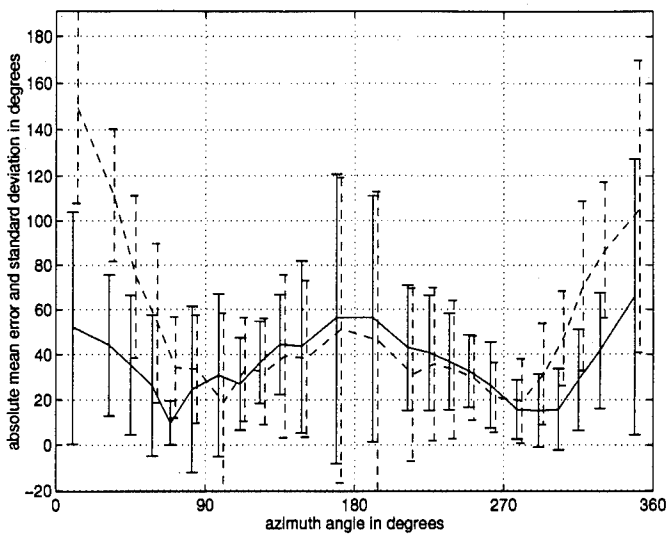


Figure 53. Absolute mean error and standard deviation for sounds with the HRTF (solid) and without the HRTF (dash). The error bars show the standard deviation of the mean error.

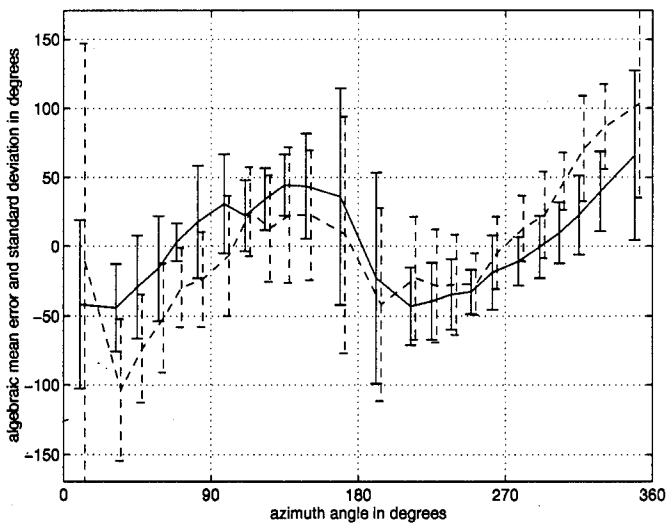


Figure 54. Algebraic mean error and standard deviation for sounds with the HRTF (solid) and without the HRTF (dash). The error bars show the standard deviation of the mean error.

direction to the error. As can be seen from the figure, the sign of the error indicates a general pulling of the perceived sound toward the ears. Figure 55 is perhaps a better view of the data. In this figure the $*$ shows the mean location of the perceived

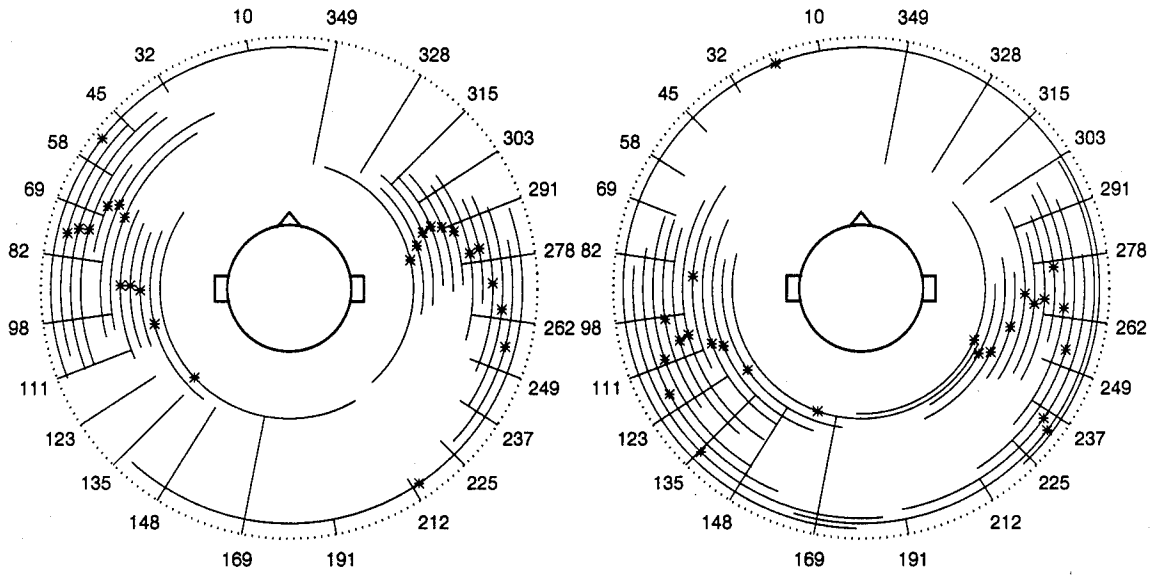


Figure 55. Mean and standard deviation for sounds with the HRTF (left) and without the HRTF (right). The arcs show the standard deviation of the mean (shown as a $*$).

sounds and the arcs show the standard deviation. As can be seen, the mean location tends to be pulled near the 90^o and 270^o locations. Figure 55 also shows how the without HRTF treatment is perceived more behind the head than in front of the head. This parallels the results of Oldfield and Parker where the pinna-filled perceptions were pulled toward the ears (18).

To further present the data in more detail, appendix A has figures 65, 66, 67 and 68 depicting all the subjects responses for each actual location. The notation used in these figures is as follows: $\times$ and solid line indicates with HRTF, $\circ$ and dashed line indicates without HRTF. The arcs show the standard deviation of the samples. The $*$ and $+$ indicate the mean with and without HRTF respectively. These are the actual responses. No corrections have been made concerning reversal.

In an effort to take into account reversals, polar histograms of the data are presented next. The polar histogram shows the frequency of subjects' responses for a specific angle of azimuth. Figure 56 shows the polar histogram of all responses for sounds with the HRTF (solid) and without the HRTF (dash). The dotted circle indicates the number of times the sounds were actual at that location. Notice subjects

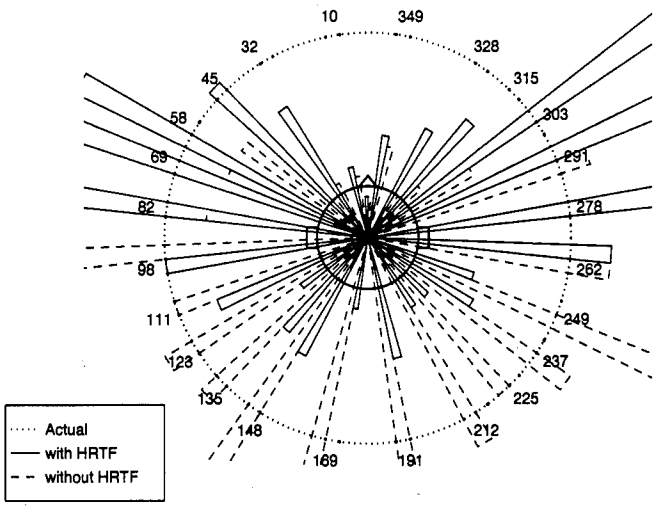


Figure 56. Polar Histogram of all responses for sounds with the HRTF (solid) and without the HRTF (dash).

heard sound at some locations more times than the sound was actual there. Recall that each subject heard 48 sounds, one with HRTF and one without HRTF for each of the 24 azimuth locations for a total of 960 sounds. This figure again indicates that subjects tended not to perceive sounds in front or behind the head. However subjects perceive less without the HRTF sounds in front of the head and more without the HRTF sound behind the head than the with HRTF sound respectively.

Tables 8 and 9 show statistics for the with HRTF and without HRTF filters. Without taking into account reversals corrections 41.25% and 34.58% of the with HRTF and without HRTF responses respectively were correct within one button ($\pm 23^o$). Recall the button layout is shown in figure 47. Figure 57 is a polar histogram of all responses correct within one button. This figure shows without HRTF correct

	with HRTF	Without HRTF
Number of samples	480	480
rmserr	0.4500	5.9583
stderr	50.7709	70.8760
correct – exact	16.88%	14.37%
correct – within one button (23°)	41.25%	34.58%

Table 8. Without reversal corrections (nor)

	with HRTF	Without HRTF
Number of samples	480	480
rmserr	0.9375	0.5458
stderr	27.6543	24.8911
Number of reversals	176	259
correct – exact	24.70%	28.75%
correct – within one button (23°)	57.29%	66.87%

Table 9. With reversal corrections (rev)

responses are greater behind the head while the with HRTF correct response are more evenly distributed front and back.

Next taking into account reversals corrections 57.29% and 66.87% of the with HRTF and without HRTF responses respectively were correct within one button ($\pm 23^o$). Figure 58 is a polar histogram of all responses with reversal corrections that are correct within one button. This figure shows without HRTF correct responses and the with HRTF correct responses are more evenly distributed front and back. The percentage of reversals with the HRTF and without the HRTF was 36.67% and 53.96% respectively. Oldfield and Parker found few reversals with the pinna unfilled and 26% with the pinna filled (17, 18). The relative increase in the number of reversals indicates that the HRTF helps in reducing the number of reversals. The differences between Oldfield and Parker and this thesis also indicate that there may be cues in the actual sound that are lost in the filtered sound. Note: Oldfield and Parker used white noise while this thesis used filtered speech.

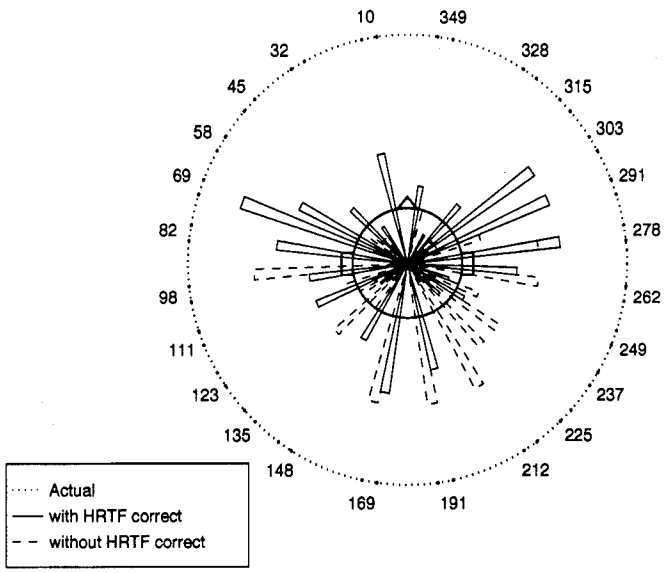


Figure 57. Polar Histogram of all responses correct within one button for sounds with the HRTF (solid) and without the HRTF (dash).

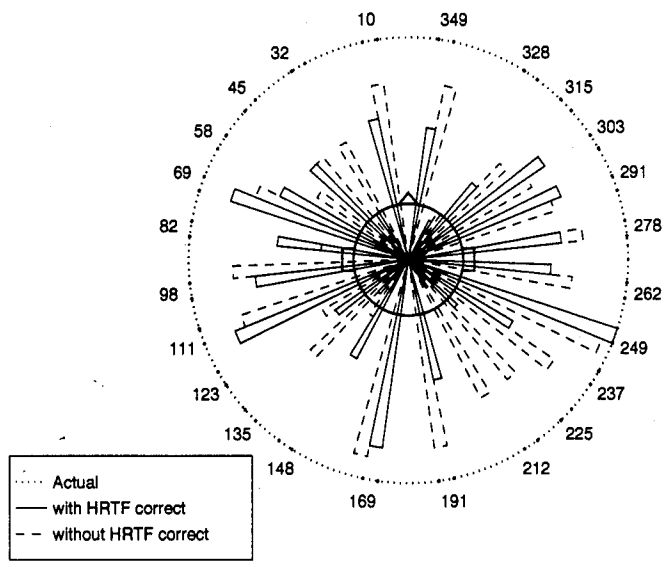


Figure 58. Polar Histogram of reversal corrected responses correct within one button for sounds with the HRTF (solid) and without the HRTF (dash).

4.3.2 Summary of Experiment One Results. Table 10 shows a summary of the ANOVA hypotheses acceptances and rejections. An "accept" means the means

Source of Variation	No Reversals Hypothesis	Reversals Hypothesis
Filter (A)	reject accept	reject accept
Azimuth (B)	reject accept	reject –
Subject (C)	reject –	reject accept
A * B	reject	accept
A * C	accept	reject
B * C	accept	accept

Table 10. Summary of Experiment One

are statistically equal, while a "reject" mean the difference between the means is statistically large. Depending on the statistical method used, we can conclude that the HRTF does provide a statistical advantage in localization accuracy over simply ITD. There is however a statistically significant interaction between the location of the sound and whether the HRTF is used. When this interaction is removed using the alternate F-Value, the statistics give mixed results of equal means for the filters and azimuth. This may mean that at certain angles of azimuth, the HRTF either provides no advantage at all or hinders localization capabilities. Adding in the calculations for reversals changes the results to where the means of the azimuth are not statistically equal. The reversal calculations will inherently reduce the error results. However comparing the number of reversals indicates an advantage of using the HRTF over no HRTF.

4.4 Experiment Two Results

The primary objective of experiment two was to determine if a neural network could be trained to learn the HRTF.

Source of Variation	Degrees of Freedom	Sum of Squares (SS)	Mean square (MS)	Calculated F Value	F at 1% level
Filter (A)	1	7578.07	7578.07	8.58	6.85,6.63
Azimuth (B)	23	287455.55	12498.07	14.15	2.03,1.79
				8.39	2.03,1.79
Subject (C)	19	39938.99	2102.05	2.38	2.19,1.88
				1.41	2.19,1.88
A * B	23	27216.78	1183.34	1.34	2.03,1.79
A * C	19	19117.75	1006.20	1.14	2.19,1.88
B * C	437	650785.39	1489.21	1.69	1.53,1.00
Error	437	385933.34	883.14		
Total	959	1418025.87			

Table 11. Analysis of Variance for "Response"

Recall, the following model for this experiment:

$$X_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \varepsilon_{ijkl}$$

Interaction effects between filters and azimuth must be tested. The null and alternate hypotheses are $H_0^{(1)}$ (no interactions between filters and azimuth) and H_1 (The means are not equal). The test statistic is F_{calc} = (Mean Square Interaction) $\div$ (Mean Square Random error) = 1.34 (as seen in table 11). Testing at an significance level of 1% ($\alpha = .01$),

$$F_{calc} = 1.34 \leq F_{.99,23,437} \approx 2.03, 1.79.$$

Hence, $H_0^{(1)}$ cannot be rejected. it can be concluded that there is no Filter * Azimuth interactions and $H_0^{(4)}$ and $H_0^{(5)}$ can be tested. Testing $H_0^{(4)}$ at an significance level of 1%,

$$F_{calc} = 8.58 > F_{.99,1,437} \approx 6.85, 6.63.$$

Testing $H_0^{(5)}$ at an significance level of 1%,

$$F_{calc} = 14.15 > F_{.99,23,437} \approx 2.03, 1.79.$$

Therefore the means of the filters nor the means of the azimuths are statistically equal.

Next completing the same tests for Filter * Subject interactions. Testing at an significance level of 1% ($\alpha = .01$),

$$F_{calc} = 1.14 \leq F_{.99,19,437} \approx 2.19, 1.88$$

shows that $H_0^{(2)}$ cannot be rejected. It can be concluded that there is no Filter * Subject interactions and $H_0^{(4)}$ and $H_0^{(6)}$ can then be tested. Testing $H_0^{(4)}$ at an significance level of 1%, $F_{calc} = 8.58 > F(.99, 1, 437) \approx 6.85, 6.63$. Testing $H_0^{(6)}$ at an significance level of 1%,

$$F_{calc} = 2.38 > F_{.99,19,437} \approx 2.19, 1.88.$$

Therefore the means of the filters nor the means of the subjects are statistically equal.

Finally, completing the same tests for Azimuth * Subject interactions. Testing at an significance level of 1% ($\alpha = .01$),

$$F_{calc} = 1.69 > F_{.99,437,437} \approx 1.53, 1.00$$

shows that $H_0^{(3)}$ can be rejected. Differences in the factors are then significant only if they are large compared with the interactions. Again the alternate F values are used. These alternate F values are the second values shown in table 6. Testing $H_0^{(5)}$ at an significance level of 1%,

$$F_{calc_{alt}} = 8.39 > F_{.99,23,437} \approx 2.03, 1.79.$$

Testing $H_0^{(6)}$ at an significance level of 1%,

$$F_{calc_{alt}} = 1.41 \leq F_{.99,19,437} \approx 2.19, 1.88.$$

Therefore the means of the azimuths are not statistically equal, however, the means of the subjects are statistically equal.

Using the reversal corrections, the following Analysis of Variance (ANOVA) (table 12) was obtained. If the reversal location error was smaller than the subjects' response then the reversal angle was used in table 12. The number of reversal corrections was 351 out of 960 or 37%.

Source of Variation	Degrees of Freedom	Sum of Squares (SS)	Mean square (MS)	Calculated F Value	F at 1% level
Filter (A)	1	1249.46	1249.46	9.23 3.94	6.85,6.63 7.88
Azimuth (B)	23	16741.34	727.88	5.38 2.29	2.03,1.79 2.78,2.70
Subject (C)	19	6288.65	330.98	2.45 –	2.19,1.88 –
A * B	23	7303.92	317.56	2.35	2.03,1.79
A * C	19	3282.8	172.78	1.28	2.19,1.88
B * C	437	89632.50	205.11	1.52	1.53,1.00
Error	437	59127.73	135.30		
Total	959	183626.43			

Table 12. Analysis of Variance for "Response" with reversal corrections

From table 12 the calculated F values indicate significant interaction between Filter * Azimuth. The other interactions Filter * Subject and Azimuth * Subject indicate no significant interaction. Comparing the within factor variance the means of filters are statistically equal; the means of the azimuths are statistically equal; and the means of the subjects are not statistically equal.

Now, the main objective of this experiment are addressed. Is there a difference in the mean response for AAMRL HRTF versus RBF HRTF binaural signals? To

determine the answer to this question, a pairwise comparison of the means for the two levels of the variable "Filter" are conducted. The following formula is used to determine the confidence intervals for the differences of two population means:

$$\bar{X}_2 - \bar{X}_1 \pm z_c \sigma_{\bar{X}_2 - \bar{X}_1} = \bar{X}_2 - \bar{X}_1 \pm z_c \sqrt{\frac{\sigma_2^2}{N_2} + \frac{\sigma_1^2}{N_1}} \tag{38}$$

where,

X_1 is the mean of the AAMRL HRTF filter error.
X_2 is the mean of the RBF HRTF filter error.
σ_1 is the standard deviation of the AAMRL HRTF filter error.
σ_2 is the standard deviation of the RBF HRTF filter error.
z_c is the 99% confidence coefficient ($z_c = 2.58$).
N_1 is the size of the AAMRL HRTF filter error.
N_2 is the size of the RBF HRTF filter error.

And since the point estimates for the means at each level are:

$$\bar{X}_1 = 38.28$$

$$\bar{X}_2 = 32.66$$

$$\bar{X}_2 - \bar{X}_1 = -5.62$$

We can state with 99 percent confidence that :

$$-12.01 < \bar{X}_2 - \bar{X}_1 < 0.77$$

And since the point estimates for the means at each level with reversal corrections are:

$$\bar{X}_1 = 17.18$$

$$\bar{X}_2 = 14.90$$

$$\bar{X}_2 - \bar{X}_1 = -2.28$$

We can state with 99 percent confidence that :

$$-4.58 < \bar{X}_2 - \bar{X}_1 < 0.02$$

The confidence interval shows a 0.77^o advantage to a 12.01^o disadvantage of using the AAMRL HRTF versus RBF HRTF. After reversal correction the advantage is reduced to 0.02^o to a disadvantage of 4.58^o. Again the correction for reversals reduced the error. The mean AAMRL HRTF error reduced from 38.28^o to 17.18^o and the mean RBF HRTF error reduced from 32.66^o to 14.90^o. The interaction of reversals is examined further in the next section.

4.4.1 Further examination of results. Figure 53 shows the absolute mean error and standard deviation for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash). The error bars show the standard deviation of the mean error. This

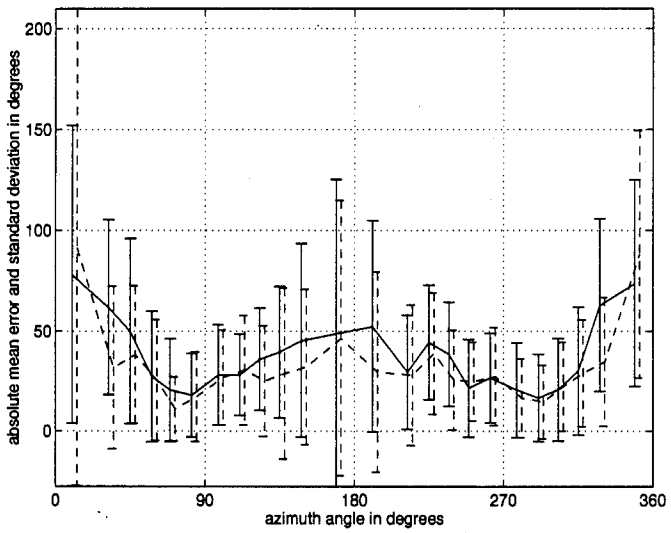


Figure 59. Absolute mean error and standard deviation for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash). The error bars show the standard deviation of the mean error.

figure reveals that AAMRL HRTF had slightly greater error than the RBF HRTF. A comparison of the standard deviations reveals no significant trends. Figure 60 shows the algebraic mean error and standard deviation for sounds with the AAMRL HRTF

(solid) and RBF HRTF (dash). The error bars show the standard deviation of the mean error. The algebraic error gives a direction to the error. As can be seen from

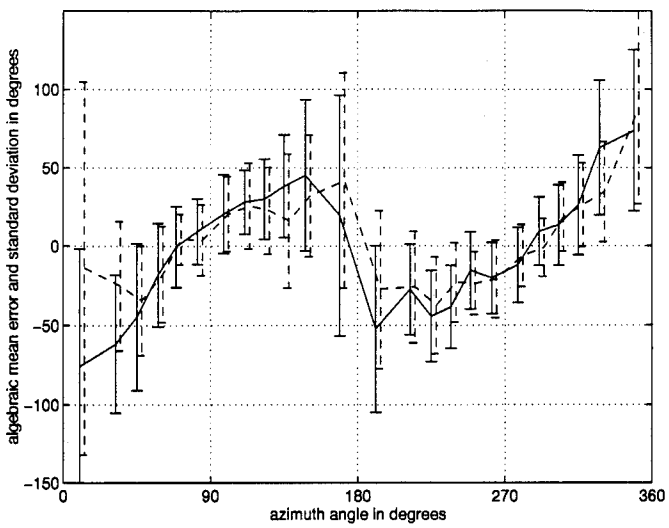


Figure 60. Algebraic mean error and standard deviation for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash). The error bars show the standard deviation of the mean error.

the figure, the sign of the error indicates a general pulling of the perceived sound toward the ears. Figure 55 is perhaps a better view of the data. In this figure the $*$ shows the mean location of the perceived sounds and the arcs show the standard deviation. As can be seen, the mean location tends to be pulled near the 90^o and 270^o locations. This parallels the results of experiment one and Oldfield and Parker where the pinna-filled perceptions were pulled toward the ears (18). Begault and Wenzle also reported subjects had a pattern of pulling toward the vertical-lateral plane (3).

To present the data in more detail, appendix B has figures 69, 70, 71 and 72 depicting all the subjects' responses for each actual location. The notation used in these figures is as follows: $\times$ and solid line indicates AAMRL HRTF, $\circ$ and dashed line indicate RBF HRTF. The arcs show the standard deviation of the samples. The $*$ and $+$ indicates the mean AAMRL and RBF HRTF respectively. These are the actual responses. No corrections have been made concerning reversal.

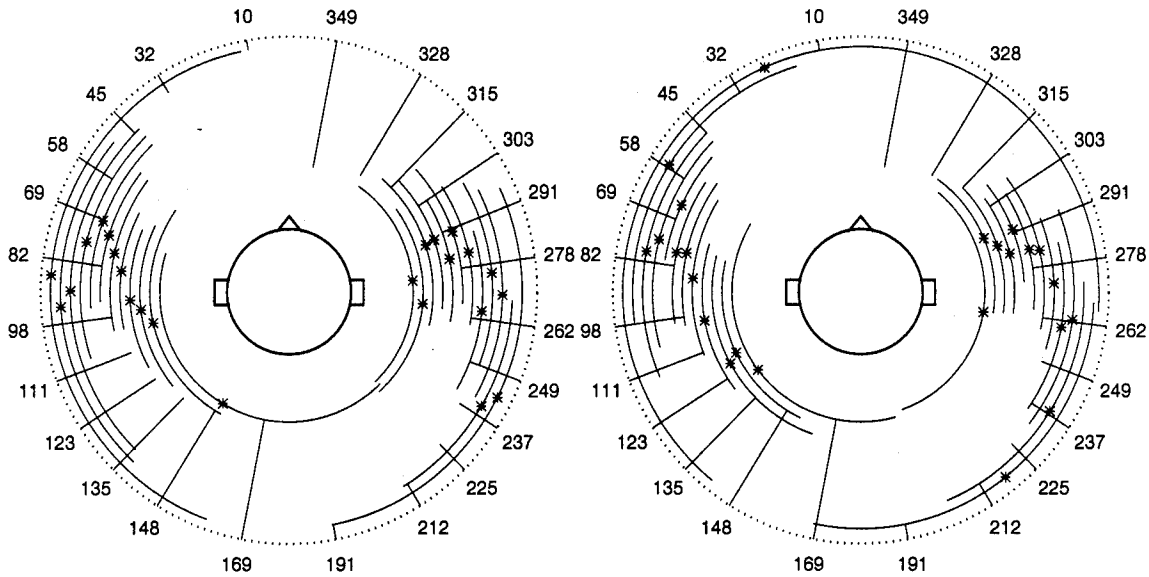


Figure 61. Mean and standard deviation for sounds with the AAMRL HRTF (left) and RBF HRTF (right). The arcs show the standard deviation of the mean (shown as a $*$).

In an effort to take into account reversals, polar histograms of the data are presented next. The polar histogram shows the frequency of subjects' responses for a specific angle of azimuth. Figure 62 shows the polar histogram of all responses for sounds with the AAMRL HRTF (solid) and the RBF HRTF (dash). The AAMRL HRTFs are label "Actual HRTF" in the figure legend. The dotted circle indicates the number times the sounds was actual at that location. Again the responses between the AAMRL HRTF and RBF HRTF are similar with the larger number of responses occurring on the sides.

Tables 13 and 14 show statistics for the AAMRL HRTF and RBF HRTF filters. Without taking into account reversals corrections 40.62% and 51.88% of the AAMRL HRTF and RBF HRTF responses respectively were correct within one button ($\pm 23^{o}$). Figure 63 is a polar histogram of all responses correct within one button. This figure shows a even distribution between filters. The front and back distribution may slightly favor the back.

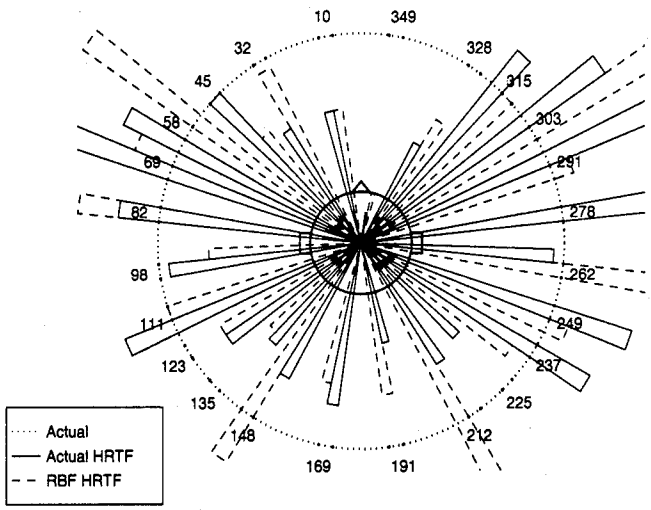


Figure 62. Polar Histogram of all responses for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash).

	AAMRL HRTF	RBF HRTF
Number of samples	480	480
rmserr	1.6354	0.0521
stderr	54.2080	50.4625
correct – exact	15.42%	22.08%
correct – within one button (23^o)	40.62%	51.88%

Table 13. Without reversal corrections (nor)

	AAMRL HRTF	RBF HRTF
Number of samples	480	480
rmserr	0.1146	0.3312
stderr	22.5449	19.9285
Number of reversals	188	163
correct – exact	25.21%	30.63%
correct – within one button (23^o)	62.71%	70.42%

Table 14. With reversal corrections (rev)

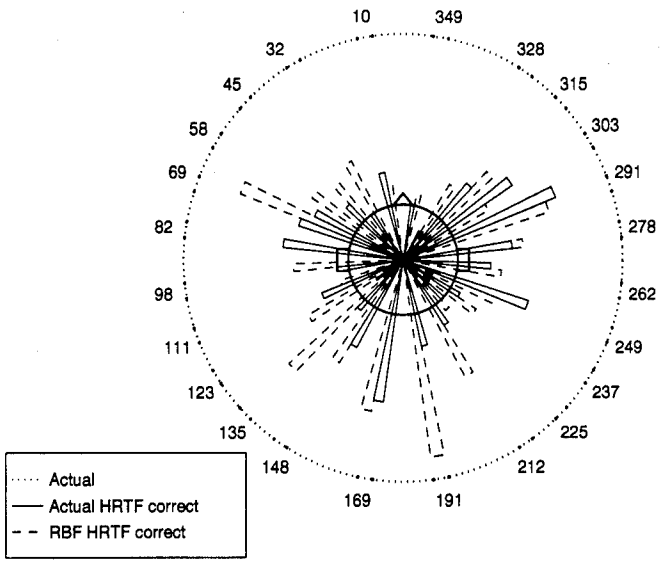


Figure 63. Polar Histogram of all responses correct within one button for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash).

Next, taking into account reversals corrections 62.71% and 70.42% of the AAMRL HRTF and RBF HRTF responses respectively were correct within one button ($\pm 23^{o}$). Figure 64 is a polar histogram of all responses with reversal corrections that are correct within one button. This figure shows there are more AAMRL HRTF correct responses and the RBF HRTF correct responses and they are more evenly distributed front and back. The percentage of reversals with the AAMRL HRTF and the RBF HRTF was 39.17% and 33.96% respectively. Begault and Wenzle reported the mean number of reversals over all the locations was 29% (3). The results of experiment one showed the AAMRL HRTF reversals to be 36.67%.

4.4.2 Summary of Experiment Two Results. Table 15 shows a summary of the ANOVA hypotheses acceptances and rejections. An "accept" means the means are statistically equal, while a "reject" mean the difference between the means is statistically large. Without considering reversals, we can conclude that the RBF HRTF provides a statistical advantage in localization accuracy over the AAMRL HRTF from which they were derived. However with reversals correction include a

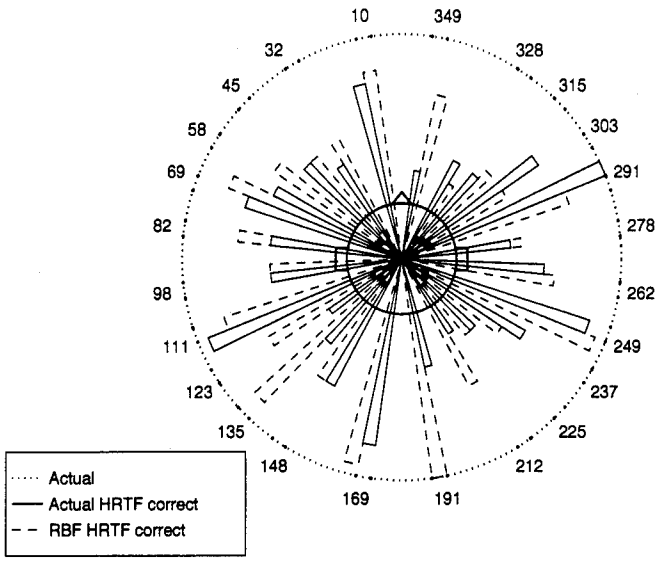


Figure 64. Polar Histogram of reversal corrected responses correct within one button for sounds with the AAMRL HRTF (solid) and RBF HRTF (dash).

Source of Variation	No Reversals Hypothesis	Reversals Hypothesis
Filter (A)	reject	reject
	–	accept
Azimuth (B)	reject	reject
	reject	accept
Subject (C)	reject	accept
	accept	–
A * B	accept	reject
A * C	accept	accept
B * C	reject	accept

Table 15. Summary of Experiment Two

statistical advantage is not indicated and the means of the two filters are statistically equal. There is also a statistically significant interaction between the location of the sound and the filter. As discussed in experiment one, this may mean that at certain angles of azimuth, the HRTF either provides no advantage at all or hinders localization capabilities. The reversal calculations will inherently reduce the error results. However, comparing the number of reversals does not indicate a large difference between the AAMRL HRTF and the RBF HRTF.

V. Conclusions and Recommendations

5.1 Summary

This thesis determined whether an artificial neural network (ANN) can approximate the Armstrong Aerospace Medical Research Laboratories (AAMRL) head related transfer functions (HRTF) data. In order to test this hypothesis, two separate tests were performed. The first test determined whether HRTF lends any support in sound localization when compared to no HRTF (Interaural Time Delay only). The second test determined whether HRTF and ANN lend the same amount of support in sound localization when compared to each other.

The following research objectives were met:

- Modified AFIT algorithms that have been used successfully on placing sounds in azimuth. The algorithms that create 3D sound with the AAMRL HRTF were changed to used the ANN HRTF instead. This allowed the creation of tests to compare the ANN HRTF with the AAMRL HRTF.
- Statistically checked to see advantage of HRTF and ITD versus ITD only.
- Statistically checked to see difference between AAMRL HRTF data and ANN HRTF data
- Investigated the program LNKmap as a neural network tool for function approximation.

5.2 Conclusions

The final conclusion is that the AAMRL HRTFs can be approximated by an artificial neural network for azimuth positions at zero degrees elevation. The point estimates for the difference in the AAMRL and RBF HRTF's mean error was only 5.62^o and 2.28^o without and with reversal corrections respectively. The point

estimates indicate that the RBF HRTF had slightly lower error than the AAMRL HRTF. However the 99% confidence interval of the mean error shows no advantage one way or the other.

To show that the HRTF was in fact contributing to sound localization, experiment one was conducted. The results indicate an advantage in localization accuracy with HRTF. There is a statistically significant interaction between the location of the sound and whether the HRTF is used. This result may indicate that for the 0^o elevations ITD has a strong influence. Comparing the numbers of reversals clearly indicates an advantage of using the HRTF over no HRTF. It is important to note that the 3D sounds presented in the experiments were the direct paths only. Smith showed an improvement in results when attenuated reflections were also included (25).

5.3 Recommendations for Future Research

The obvious next step is to expand the neural network training the other elevations of the AAMRL data. The AAMRL data has HRTF responses for elevation $\pm 90^o$. Oldfield and Parker found that the acuity of sound localization varied with azimuth and elevation (17). The ANN HRTF would have to be tested in these non-zero degrees elevation regions.

The AAMRL HRTF 0^o elevation data in this thesis was spaced approximately 23^o apart. Future research could investigate the interpolations accuracy of the ANN against the AAMRL HRTF data spaced 1^o apart at 0^o elevation.

Replacing the AAMRL HRTF samples by an artificial neural network approximator may also provide insight of underlining functions within the HRTF. Since subjects where able to localize with the smoothed ANN HRTF, simpler filter designs may be implemented.

Implementing the ANN HRTF in hardware may provide an alternate approach to real-time implementation other than the DIRAD and Convolvotron.

Appendix A. Experiment 1 Data: With and Without HRTF

Actual	With HRTF	Without HRTF
10^o	148^o	169^o
32^o	32^o	148^o
45^o	123^o	148^o
58^o	58^o	148^o
69^o	69^o	82^o
82^o	58^o	111^o
98^o	82^o	123^o
111^o	82^o	82^o
123^o	135^o	135^o
135^o	69^o	135^o
148^o	148^o	123^o
169^o	148^o	169^o
191^o	191^o	169^o
212^o	249^o	212^o
225^o	262^o	212^o
237^o	262^o	291^o
249^o	278^o	278^o
262^o	237^o	278^o
278^o	303^o	249^o
291^o	315^o	237^o
303^o	291^o	249^o
315^o	291^o	212^o
328^o	278^o	249^o
349^o	349^o	191^o

Actual	With HRTF	Without HRTF
10^o	10^o	191^o
32^o	148^o	169^o
45^o	148^o	169^o
58^o	82^o	148^o
69^o	45^o	98^o
82^o	69^o	135^o
98^o	69^o	111^o
111^o	98^o	123^o
123^o	135^o	148^o
135^o	111^o	169^o
148^o	111^o	135^o
169^o	169^o	169^o
191^o	212^o	191^o
212^o	225^o	212^o
225^o	225^o	225^o
237^o	303^o	225^o
249^o	278^o	315^o
262^o	262^o	262^o
278^o	262^o	303^o
291^o	278^o	249^o
303^o	291^o	291^o
315^o	349^o	212^o
328^o	225^o	278^o
349^o	328^o	191^o

Actual	With HRTF	Without HRTF
10^o	32^o	191^o
32^o	45^o	135^o
45^o	82^o	148^o
58^o	45^o	148^o
69^o	69^o	135^o
82^o	45^o	82^o
98^o	45^o	98^o
111^o	82^o	98^o
123^o	82^o	148^o
135^o	82^o	123^o
148^o	135^o	169^o
169^o	169^o	169^o
191^o	349^o	191^o
212^o	212^o	212^o
225^o	225^o	262^o
237^o	278^o	249^o
249^o	278^o	291^o
262^o	315^o	262^o
278^o	328^o	278^o
291^o	315^o	291^o
303^o	315^o	262^o
315^o	328^o	225^o
328^o	303^o	262^o
349^o	349^o	0^o

Actual	With HRTF	Without HRTF
10^o	10^o	191^o
32^o	45^o	169^o
45^o	32^o	135^o
58^o	45^o	135^o
69^o	58^o	135^o
82^o	45^o	135^o
98^o	69^o	98^o
111^o	69^o	111^o
123^o	82^o	148^o
135^o	82^o	148^o
148^o	148^o	148^o
169^o	191^o	169^o
191^o	212^o	191^o
212^o	278^o	212^o
225^o	278^o	225^o
237^o	278^o	249^o
249^o	303^o	278^o
262^o	278^o	291^o
278^o	291^o	278^o
291^o	303^o	225^o
303^o	315^o	249^o
315^o	315^o	225^o
328^o	303^o	237^o
349^o	349^o	191^o

Actual	With HRTF	Without HRTF
10^o	82^o	148^o
32^o	82^o	148^o
45^o	10^o	123^o
58^o	32^o	98^o
69^o	69^o	111^o
82^o	278^o	69^o
98^o	45^o	69^o
111^o	98^o	82^o
123^o	0^o	123^o
135^o	111^o	123^o
148^o	82^o	0^o
169^o	111^o	148^o
191^o	148^o	169^o
212^o	237^o	278^o
225^o	0^o	303^o
237^o	315^o	291^o
249^o	303^o	303^o
262^o	262^o	249^o
278^o	291^o	225^o
291^o	278^o	328^o
303^o	278^o	212^o
315^o	291^o	212^o
328^o	328^o	212^o
349^o	191^o	169^o

Actual	With HRTF	Without HRTF
10^o	349^o	169^o
32^o	123^o	315^o
45^o	69^o	148^o
58^o	58^o	58^o
69^o	45^o	58^o
82^o	82^o	135^o
98^o	291^o	69^o
111^o	111^o	123^o
123^o	98^o	148^o
135^o	111^o	69^o
148^o	32^o	148^o
169^o	169^o	32^o
191^o	191^o	349^o
212^o	249^o	212^o
225^o	303^o	212^o
237^o	278^o	191^o
249^o	278^o	278^o
262^o	315^o	237^o
278^o	315^o	212^o
291^o	303^o	291^o
303^o	0^o	225^o
315^o	328^o	191^o
328^o	291^o	148^o
349^o	328^o	10^o

Actual	With HRTF	Without HRTF
10^o	32^o	45^o
32^o	32^o	123^o
45^o	69^o	111^o
58^o	82^o	98^o
69^o	58^o	82^o
82^o	32^o	98^o
98^o	58^o	82^o
111^o	45^o	69^o
123^o	82^o	82^o
135^o	98^o	98^o
148^o	45^o	111^o
169^o	32^o	169^o
191^o	191^o	191^o
212^o	291^o	303^o
225^o	291^o	328^o
237^o	278^o	315^o
249^o	303^o	249^o
262^o	278^o	278^o
278^o	303^o	278^o
291^o	315^o	278^o
303^o	315^o	291^o
315^o	291^o	303^o
328^o	262^o	249^o
349^o	303^o	212^o

Actual	With HRTF	Without HRTF
10^o	32^o	169^o
32^o	82^o	148^o
45^o	69^o	123^o
58^o	58^o	169^o
69^o	58^o	123^o
82^o	69^o	123^o
98^o	82^o	123^o
111^o	98^o	82^o
123^o	58^o	111^o
135^o	98^o	148^o
148^o	135^o	98^o
169^o	148^o	191^o
191^o	58^o	191^o
212^o	237^o	191^o
225^o	262^o	237^o
237^o	249^o	303^o
249^o	262^o	291^o
262^o	291^o	278^o
278^o	278^o	278^o
291^o	291^o	249^o
303^o	315^o	291^o
315^o	237^o	237^o
328^o	278^o	262^o
349^o	191^o	212^o

Actual	With HRTF	Without HRTF
10^o	32^o	191^o
32^o	123^o	169^o
45^o	32^o	135^o
58^o	45^o	148^o
69^o	45^o	111^o
82^o	58^o	135^o
98^o	69^o	82^o
111^o	82^o	123^o
123^o	82^o	135^o
135^o	135^o	148^o
148^o	135^o	169^o
169^o	10^o	169^o
191^o	212^o	191^o
212^o	237^o	225^o
225^o	237^o	225^o
237^o	303^o	237^o
249^o	278^o	262^o
262^o	278^o	249^o
278^o	291^o	278^o
291^o	291^o	225^o
303^o	303^o	262^o
315^o	237^o	262^o
328^o	303^o	225^o
349^o	349^o	191^o

Actual	With HRTF	Without HRTF
10^o	32^o	169^o
32^o	82^o	98^o
45^o	58^o	98^o
58^o	191^o	98^o
69^o	58^o	98^o
82^o	69^o	123^o
98^o	58^o	98^o
111^o	123^o	69^o
123^o	58^o	98^o
135^o	98^o	98^o
148^o	58^o	98^o
169^o	111^o	123^o
191^o	191^o	349^o
212^o	237^o	225^o
225^o	262^o	0^o
237^o	262^o	237^o
249^o	303^o	249^o
262^o	291^o	249^o
278^o	303^o	262^o
291^o	237^o	315^o
303^o	303^o	262^o
315^o	278^o	291^o
328^o	349^o	225^o
349^o	249^o	212^o

Actual	With HRTF	Without HRTF
10^o	328^o	169^o
32^o	69^o	148^o
45^o	123^o	169^o
58^o	82^o	123^o
69^o	69^o	98^o
82^o	82^o	135^o
98^o	98^o	111^o
111^o	69^o	45^o
123^o	82^o	148^o
135^o	58^o	123^o
148^o	58^o	69^o
169^o	169^o	191^o
191^o	278^o	191^o
212^o	278^o	212^o
225^o	291^o	191^o
237^o	278^o	349^o
249^o	291^o	303^o
262^o	328^o	315^o
278^o	303^o	262^o
291^o	303^o	237^o
303^o	291^o	262^o
315^o	303^o	191^o
328^o	303^o	249^o
349^o	303^o	237^o

Actual	With HRTF	Without HRTF
10^o	58^o	262^o
32^o	82^o	111^o
45^o	69^o	111^o
58^o	69^o	111^o
69^o	98^o	98^o
82^o	98^o	82^o
98^o	82^o	98^o
111^o	98^o	98^o
123^o	69^o	98^o
135^o	69^o	98^o
148^o	69^o	0^o
169^o	349^o	349^o
191^o	291^o	249^o
212^o	278^o	278^o
225^o	262^o	291^o
237^o	249^o	278^o
249^o	278^o	278^o
262^o	278^o	262^o
278^o	278^o	262^o
291^o	291^o	278^o
303^o	262^o	262^o
315^o	278^o	278^o
328^o	278^o	278^o
349^o	262^o	278^o

Actual	With HRTF	Without HRTF
10^o	169^o	169^o
32^o	58^o	191^o
45^o	58^o	58^o
58^o	69^o	58^o
69^o	69^o	111^o
82^o	82^o	58^o
98^o	98^o	98^o
111^o	58^o	69^o
123^o	82^o	45^o
135^o	58^o	58^o
148^o	111^o	32^o
169^o	32^o	10^o
191^o	32^o	10^o
212^o	328^o	349^o
225^o	303^o	303^o
237^o	278^o	262^o
249^o	303^o	291^o
262^o	291^o	249^o
278^o	278^o	291^o
291^o	291^o	262^o
303^o	291^o	249^o
315^o	278^o	303^o
328^o	303^o	237^o
349^o	191^o	328^o

Actual	With HRTF	Without HRTF
10^o	135^o	191^o
32^o	69^o	148^o
45^o	45^o	123^o
58^o	45^o	148^o
69^o	69^o	58^o
82^o	82^o	58^o
98^o	58^o	98^o
111^o	111^o	69^o
123^o	82^o	58^o
135^o	69^o	169^o
148^o	111^o	111^o
169^o	148^o	191^o
191^o	249^o	191^o
212^o	237^o	191^o
225^o	278^o	249^o
237^o	291^o	249^o
249^o	291^o	291^o
262^o	278^o	291^o
278^o	291^o	291^o
291^o	278^o	262^o
303^o	303^o	278^o
315^o	303^o	225^o
328^o	291^o	278^o
349^o	303^o	212^o

Actual	With HRTF	Without HRTF
10^o	10^o	169^o
32^o	58^o	169^o
45^o	45^o	32^o
58^o	45^o	58^o
69^o	69^o	45^o
82^o	82^o	58^o
98^o	82^o	0^o
111^o	111^o	69^o
123^o	98^o	45^o
135^o	98^o	169^o
148^o	135^o	111^o
169^o	169^o	169^o
191^o	212^o	191^o
212^o	237^o	191^o
225^o	278^o	328^o
237^o	303^o	249^o
249^o	262^o	291^o
262^o	278^o	291^o
278^o	291^o	291^o
291^o	303^o	291^o
303^o	315^o	249^o
315^o	315^o	212^o
328^o	328^o	237^o
349^o	349^o	349^o

Actual	With HRTF	Without HRTF
10^o	58^o	169^o
32^o	98^o	135^o
45^o	123^o	58^o
58^o	135^o	135^o
69^o	69^o	69^o
82^o	69^o	148^o
98^o	98^o	98^o
111^o	45^o	135^o
123^o	98^o	58^o
135^o	82^o	32^o
148^o	0^o	69^o
169^o	148^o	148^o
191^o	303^o	249^o
212^o	262^o	237^o
225^o	291^o	237^o
237^o	249^o	249^o
249^o	278^o	262^o
262^o	212^o	225^o
278^o	262^o	237^o
291^o	291^o	225^o
303^o	225^o	237^o
315^o	315^o	303^o
328^o	262^o	237^o
349^o	291^o	237^o

Actual	With HRTF	Without HRTF
10^o	111^o	191^o
32^o	69^o	123^o
45^o	123^o	111^o
58^o	111^o	111^o
69^o	82^o	98^o
82^o	69^o	82^o
98^o	82^o	82^o
111^o	82^o	69^o
123^o	82^o	98^o
135^o	82^o	148^o
148^o	123^o	123^o
169^o	148^o	135^o
191^o	262^o	191^o
212^o	237^o	237^o
225^o	249^o	278^o
237^o	249^o	249^o
249^o	262^o	278^o
262^o	262^o	278^o
278^o	262^o	278^o
291^o	262^o	249^o
303^o	303^o	249^o
315^o	291^o	249^o
328^o	262^o	225^o
349^o	291^o	169^o

Actual	With HRTF	Without HRTF
10^o	10^o	212^o
32^o	111^o	191^o
45^o	45^o	169^o
58^o	45^o	32^o
69^o	82^o	148^o
82^o	135^o	169^o
98^o	98^o	278^o
111^o	148^o	10^o
123^o	111^o	169^o
135^o	135^o	10^o
148^o	58^o	262^o
169^o	10^o	10^o
191^o	191^o	349^o
212^o	237^o	225^o
225^o	212^o	212^o
237^o	225^o	291^o
249^o	249^o	237^o
262^o	303^o	315^o
278^o	278^o	249^o
291^o	225^o	315^o
303^o	278^o	249^o
315^o	278^o	212^o
328^o	278^o	249^o
349^o	191^o	349^o

Actual	With HRTF	Without HRTF
10^o	148^o	169^o
32^o	58^o	148^o
45^o	82^o	135^o
58^o	98^o	111^o
69^o	58^o	135^o
82^o	58^o	111^o
98^o	69^o	82^o
111^o	69^o	135^o
123^o	69^o	123^o
135^o	82^o	148^o
148^o	135^o	111^o
169^o	169^o	148^o
191^o	212^o	191^o
212^o	278^o	191^o
225^o	278^o	237^o
237^o	225^o	237^o
249^o	278^o	237^o
262^o	303^o	225^o
278^o	291^o	249^o
291^o	291^o	278^o
303^o	291^o	237^o
315^o	262^o	262^o
328^o	249^o	212^o
349^o	191^o	191^o

Actual	With HRTF	Without HRTF
10^o	328^o	328^o
32^o	58^o	82^o
45^o	82^o	69^o
58^o	82^o	58^o
69^o	82^o	82^o
82^o	82^o	69^o
98^o	69^o	98^o
111^o	98^o	58^o
123^o	69^o	82^o
135^o	82^o	32^o
148^o	111^o	123^o
169^o	58^o	349^o
191^o	291^o	303^o
212^o	278^o	303^o
225^o	278^o	278^o
237^o	278^o	291^o
249^o	278^o	262^o
262^o	278^o	249^o
278^o	278^o	249^o
291^o	278^o	278^o
303^o	278^o	249^o
315^o	278^o	278^o
328^o	291^o	278^o
349^o	303^o	278^o

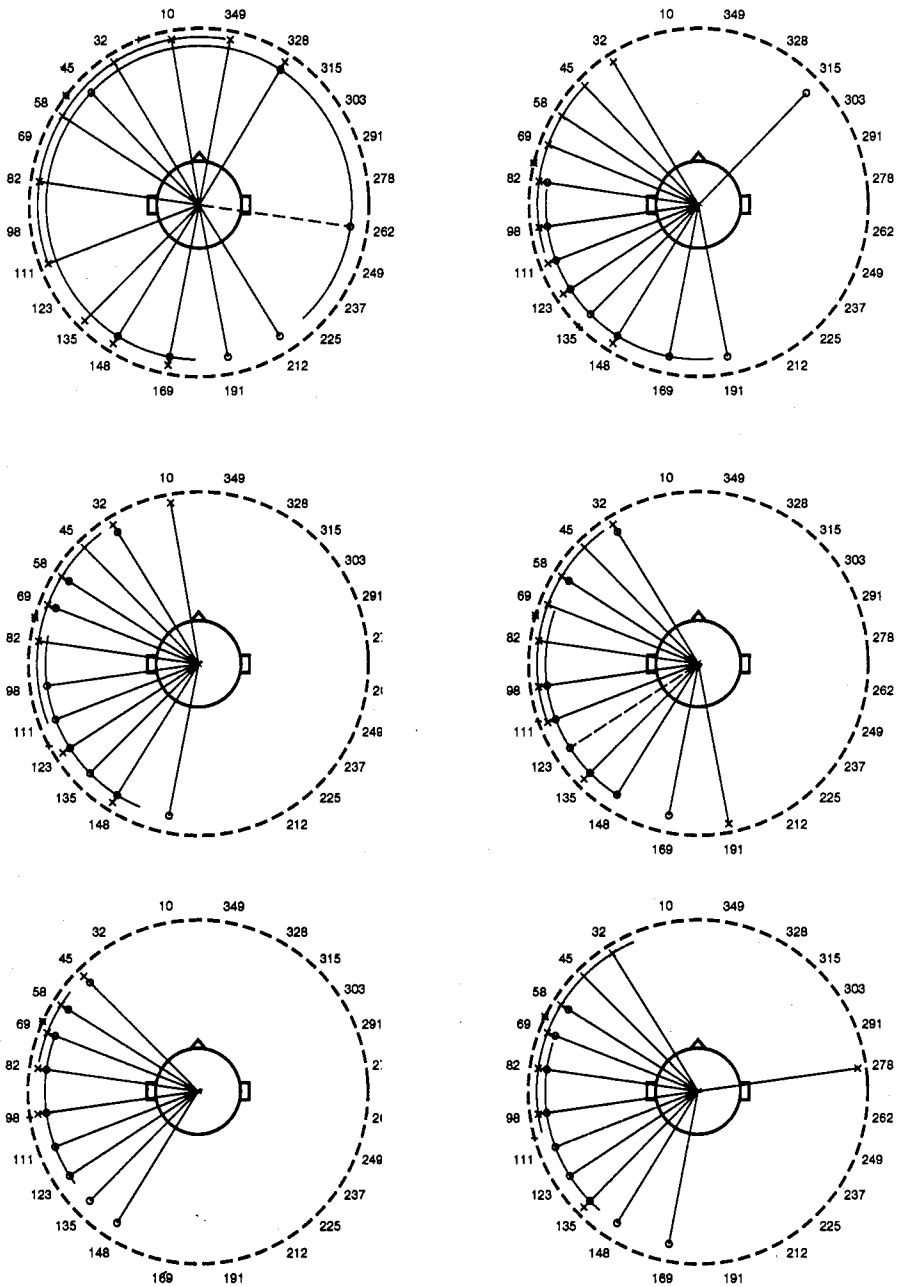


Figure 65. All subjects' responses for actual locations, 10^o, 32^o, 45^o, 58^o, 69^o and 82^o shown left to right, top to bottom respectively. Notation: $\times$ and solid line indicates with HRTF, $\circ$ and dashed line indicate without HRTF. The arcs show the standard deviation of the samples. The $\star$ and $+$ indicates the mean with and without HRTF respectively.

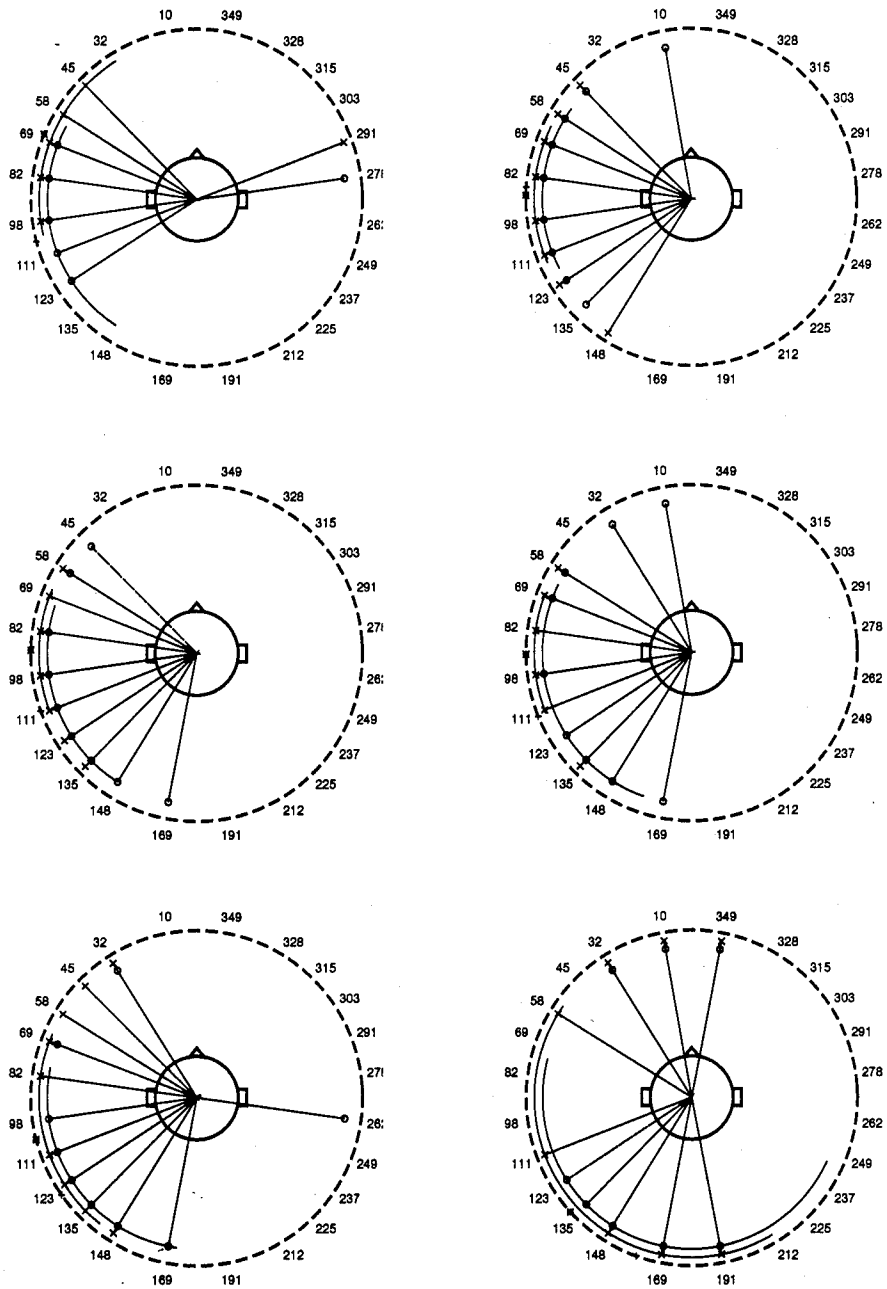


Figure 66. All subjects' responses for actual locations, 98°, 111°, 123°, 135°, 148° and 169° shown left to right, top to bottom respectively. Notation is same as above.

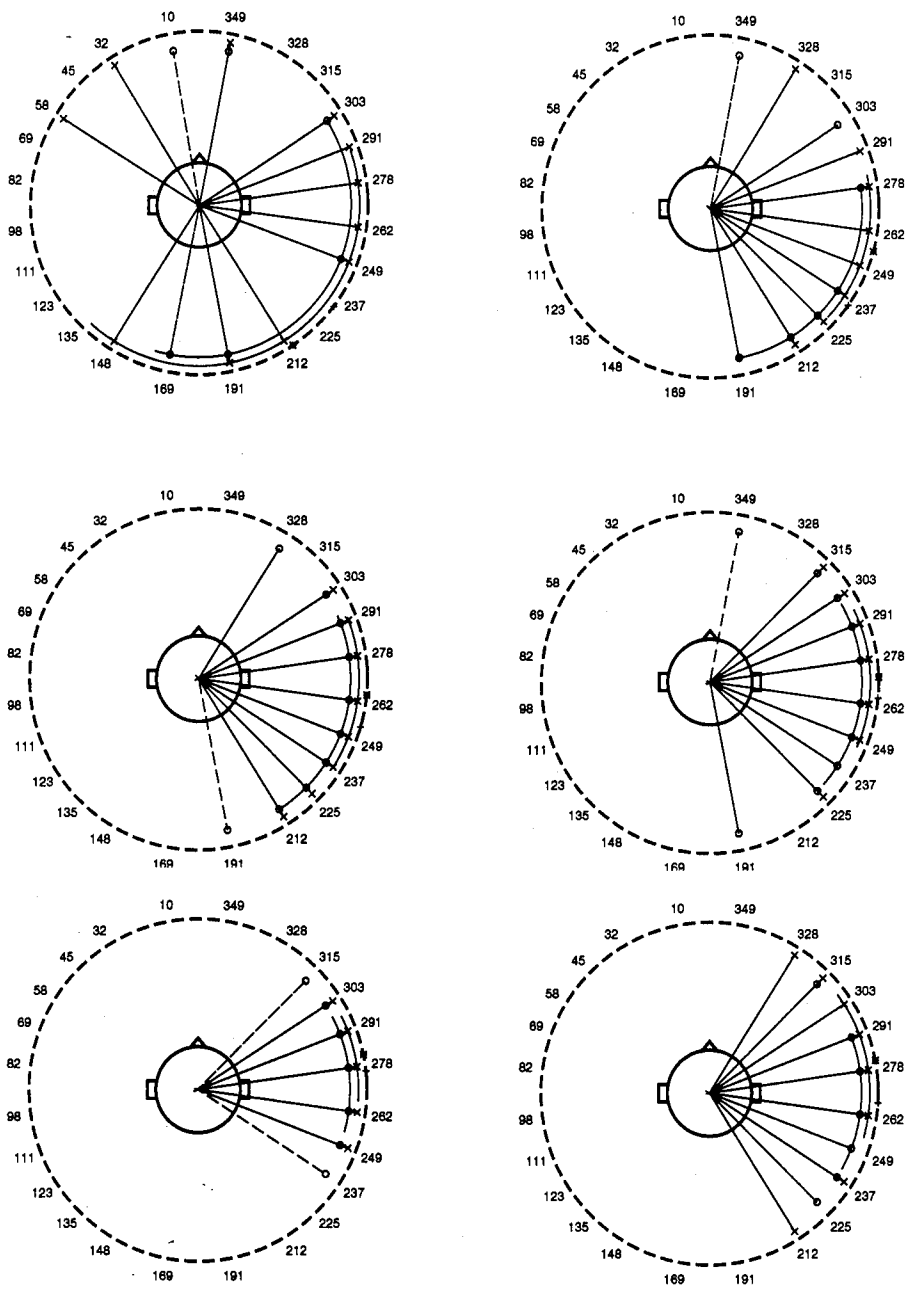


Figure 67. All subjects' responses for actual locations, 191°, 212°, 225°, 237°, 249° and 262° shown left to right, top to bottom respectively. Notation is same as above.

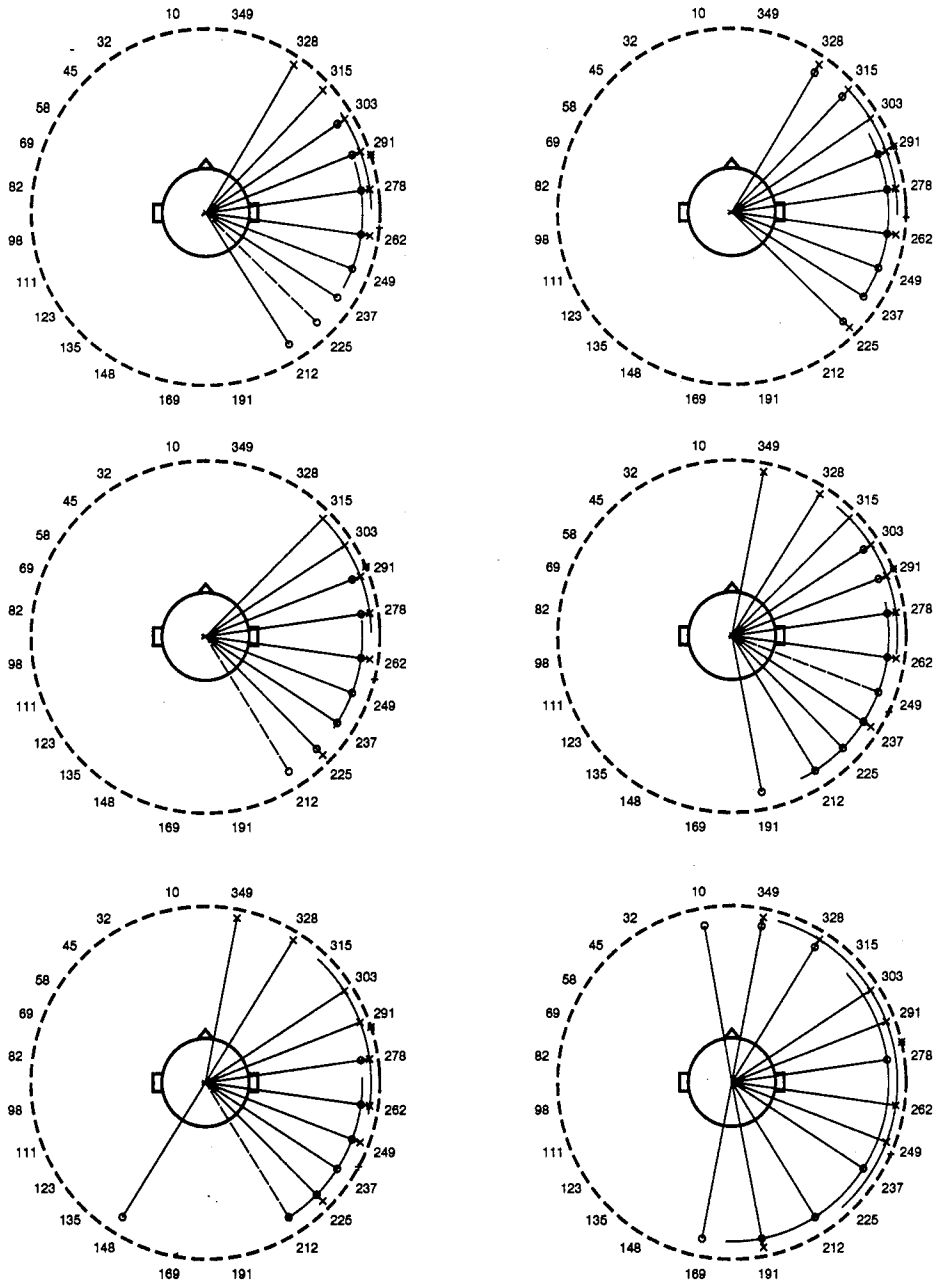


Figure 68. All subjects' responses for actual locations, 278°, 291°, 303°, 315°, 328° and 349° shown left to right, top to bottom respectively. Notation is same as above.

Appendix B. Experiment 2 Data: AAMRL HRTF and RBF HRTF

Actual	AAMRL HRTF	RBF HRTF
10^o	111^o	148^o
32^o	69^o	58^o
45^o	69^o	82^o
58^o	45^o	58^o
69^o	58^o	45^o
82^o	58^o	98^o
98^o	45^o	98^o
111^o	82^o	69^o
123^o	98^o	98^o
135^o	111^o	123^o
148^o	148^o	135^o
169^o	169^o	169^o
191^o	191^o	212^o
212^o	262^o	249^o
225^o	303^o	278^o
237^o	262^o	262^o
249^o	291^o	291^o
262^o	303^o	315^o
278^o	291^o	303^o
291^o	291^o	291^o
303^o	262^o	262^o
315^o	249^o	249^o
328^o	328^o	328^o
349^o	225^o	191^o

Actual	AAMRL HRTF	RBF HRTF
10^o	10^o	10^o
32^o	82^o	58^o
45^o	69^o	111^o
58^o	69^o	82^o
69^o	82^o	69^o
82^o	111^o	82^o
98^o	98^o	82^o
111^o	111^o	111^o
123^o	98^o	123^o
135^o	69^o	135^o
148^o	10^o	98^o
169^o	32^o	10^o
191^o	237^o	212^o
212^o	249^o	212^o
225^o	262^o	249^o
237^o	249^o	249^o
249^o	237^o	262^o
262^o	278^o	262^o
278^o	262^o	262^o
291^o	262^o	278^o
303^o	278^o	262^o
315^o	249^o	249^o
328^o	212^o	262^o
349^o	237^o	303^o

Actual	AAMRL HRTF	RBF HRTF
10^o	10^o	191^o
32^o	69^o	32^o
45^o	148^o	123^o
58^o	135^o	98^o
69^o	58^o	58^o
82^o	69^o	111^o
98^o	45^o	69^o
111^o	82^o	58^o
123^o	82^o	123^o
135^o	82^o	148^o
148^o	148^o	148^o
169^o	191^o	169^o
191^o	278^o	212^o
212^o	212^o	212^o
225^o	278^o	212^o
237^o	249^o	249^o
249^o	249^o	262^o
262^o	291^o	303^o
278^o	303^o	303^o
291^o	303^o	262^o
303^o	315^o	262^o
315^o	291^o	291^o
328^o	225^o	303^o
349^o	191^o	212^o

Actual	AAMRL HRTF	RBF HRTF
10^o	10^o	349^o
32^o	45^o	32^o
45^o	32^o	111^o
58^o	32^o	58^o
69^o	69^o	58^o
82^o	69^o	45^o
98^o	69^o	82^o
111^o	111^o	45^o
123^o	135^o	123^o
135^o	82^o	148^o
148^o	123^o	148^o
169^o	349^o	169^o
191^o	225^o	191^o
212^o	191^o	225^o
225^o	278^o	237^o
237^o	291^o	262^o
249^o	291^o	262^o
262^o	278^o	278^o
278^o	278^o	303^o
291^o	291^o	303^o
303^o	315^o	303^o
315^o	328^o	303^o
328^o	291^o	315^o
349^o	328^o	349^o

Actual	AAMRL HRTF	RBF HRTF
10^o	169^o	32^o
32^o	123^o	10^o
45^o	169^o	58^o
58^o	58^o	98^o
69^o	98^o	69^o
82^o	82^o	69^o
98^o	111^o	82^o
111^o	98^o	98^o
123^o	98^o	69^o
135^o	135^o	148^o
148^o	135^o	148^o
169^o	169^o	169^o
191^o	303^o	212^o
212^o	212^o	212^o
225^o	225^o	291^o
237^o	237^o	237^o
249^o	278^o	278^o
262^o	278^o	291^o
278^o	278^o	291^o
291^o	303^o	278^o
303^o	291^o	249^o
315^o	291^o	328^o
328^o	225^o	237^o
349^o	191^o	212^o

Actual	AAMRL HRTF	RBF HRTF
10^o	169^o	191^o
32^o	111^o	32^o
45^o	111^o	32^o
58^o	45^o	45^o
69^o	45^o	69^o
82^o	69^o	82^o
98^o	58^o	58^o
111^o	98^o	82^o
123^o	82^o	111^o
135^o	123^o	169^o
148^o	135^o	135^o
169^o	169^o	169^o
191^o	212^o	191^o
212^o	212^o	212^o
225^o	249^o	249^o
237^o	291^o	262^o
249^o	249^o	278^o
262^o	278^o	278^o
278^o	291^o	278^o
291^o	278^o	315^o
303^o	315^o	303^o
315^o	315^o	328^o
328^o	328^o	328^o
349^o	349^o	328^o

Actual	AAMRL HRTF	RBF HRTF
10^o	10^o	10^o
32^o	169^o	32^o
45^o	45^o	69^o
58^o	58^o	69^o
69^o	58^o	82^o
82^o	82^o	69^o
98^o	69^o	69^o
111^o	69^o	58^o
123^o	98^o	111^o
135^o	69^o	135^o
148^o	32^o	148^o
169^o	169^o	148^o
191^o	212^o	212^o
212^o	249^o	262^o
225^o	262^o	249^o
237^o	278^o	262^o
249^o	291^o	278^o
262^o	262^o	291^o
278^o	303^o	278^o
291^o	291^o	278^o
303^o	291^o	278^o
315^o	291^o	303^o
328^o	237^o	278^o
349^o	249^o	249^o

Actual	AAMRL HRTF	RBF HRTF
10^o	10^o	191^o
32^o	148^o	45^o
45^o	169^o	58^o
58^o	148^o	32^o
69^o	32^o	58^o
82^o	82^o	82^o
98^o	82^o	45^o
111^o	82^o	82^o
123^o	169^o	135^o
135^o	32^o	32^o
148^o	111^o	82^o
169^o	191^o	10^o
191^o	191^o	349^o
212^o	262^o	262^o
225^o	303^o	237^o
237^o	278^o	278^o
249^o	225^o	262^o
262^o	237^o	249^o
278^o	315^o	291^o
291^o	278^o	278^o
303^o	278^o	278^o
315^o	328^o	303^o
328^o	291^o	291^o
349^o	249^o	349^o

Actual	AAMRL HRTF	RBF HRTF
10^o	169^o	328^o
32^o	45^o	10^o
45^o	32^o	32^o
58^o	45^o	135^o
69^o	32^o	32^o
82^o	58^o	58^o
98^o	69^o	45^o
111^o	58^o	82^o
123^o	58^o	45^o
135^o	45^o	148^o
148^o	111^o	69^o
169^o	10^o	191^o
191^o	249^o	191^o
212^o	249^o	291^o
225^o	237^o	315^o
237^o	303^o	262^o
249^o	315^o	249^o
262^o	249^o	278^o
278^o	303^o	278^o
291^o	262^o	315^o
303^o	303^o	278^o
315^o	291^o	315^o
328^o	303^o	291^o
349^o	315^o	212^o

Actual	AAMRL HRTF	RBF HRTF
10^o	148^o	148^o
32^o	58^o	58^o
45^o	32^o	111^o
58^o	32^o	69^o
69^o	98^o	69^o
82^o	45^o	32^o
98^o	69^o	82^o
111^o	69^o	148^o
123^o	98^o	111^o
135^o	123^o	148^o
148^o	123^o	123^o
169^o	148^o	169^o
191^o	249^o	169^o
212^o	237^o	212^o
225^o	237^o	328^o
237^o	249^o	212^o
249^o	262^o	291^o
262^o	237^o	212^o
278^o	262^o	303^o
291^o	303^o	303^o
303^o	249^o	262^o
315^o	315^o	262^o
328^o	212^o	278^o
349^o	315^o	212^o

Actual	AAMRL HRTF	RBF HRTF
10^o	58^o	169^o
32^o	123^o	135^o
45^o	58^o	98^o
58^o	58^o	45^o
69^o	98^o	69^o
82^o	58^o	69^o
98^o	45^o	111^o
111^o	111^o	82^o
123^o	98^o	58^o
135^o	148^o	148^o
148^o	123^o	135^o
169^o	69^o	10^o
191^o	191^o	212^o
212^o	225^o	191^o
225^o	291^o	249^o
237^o	237^o	262^o
249^o	262^o	291^o
262^o	291^o	278^o
278^o	225^o	237^o
291^o	237^o	303^o
303^o	315^o	303^o
315^o	303^o	249^o
328^o	278^o	303^o
349^o	328^o	328^o

Actual	AAMRL HRTF	RBF HRTF
10^o	10^o	169^o
32^o	45^o	32^o
45^o	123^o	148^o
58^o	82^o	111^o
69^o	123^o	69^o
82^o	123^o	111^o
98^o	111^o	111^o
111^o	98^o	111^o
123^o	69^o	135^o
135^o	111^o	135^o
148^o	135^o	148^o
169^o	148^o	32^o
191^o	249^o	212^o
212^o	291^o	237^o
225^o	291^o	303^o
237^o	315^o	315^o
249^o	237^o	315^o
262^o	315^o	315^o
278^o	315^o	303^o
291^o	303^o	315^o
303^o	315^o	315^o
315^o	328^o	303^o
328^o	212^o	328^o
349^o	291^o	328^o

Actual	AAMRL HRTF	RBF HRTF
10^o	349^o	10^o
32^o	148^o	148^o
45^o	123^o	58^o
58^o	111^o	82^o
69^o	58^o	69^o
82^o	69^o	111^o
98^o	123^o	123^o
111^o	69^o	123^o
123^o	69^o	123^o
135^o	135^o	135^o
148^o	148^o	148^o
169^o	169^o	148^o
191^o	225^o	191^o
212^o	237^o	212^o
225^o	315^o	237^o
237^o	315^o	249^o
249^o	303^o	237^o
262^o	303^o	291^o
278^o	328^o	303^o
291^o	303^o	315^o
303^o	237^o	315^o
315^o	303^o	315^o
328^o	225^o	328^o
349^o	212^o	315^o

Actual	AAMRL HRTF	RBF HRTF
10^o	169^o	328^o
32^o	148^o	10^o
45^o	135^o	58^o
58^o	111^o	123^o
69^o	69^o	58^o
82^o	69^o	58^o
98^o	58^o	111^o
111^o	58^o	82^o
123^o	111^o	82^o
135^o	123^o	58^o
148^o	10^o	58^o
169^o	212^o	191^o
191^o	249^o	212^o
212^o	237^o	303^o
225^o	249^o	278^o
237^o	291^o	262^o
249^o	237^o	278^o
262^o	303^o	291^o
278^o	278^o	278^o
291^o	291^o	278^o
303^o	303^o	315^o
315^o	291^o	303^o
328^o	249^o	315^o
349^o	303^o	212^o

Actual	AAMRL HRTF	RBF HRTF
10^o	45^o	237^o
32^o	32^o	58^o
45^o	58^o	45^o
58^o	111^o	58^o
69^o	111^o	69^o
82^o	98^o	69^o
98^o	82^o	58^o
111^o	98^o	111^o
123^o	82^o	98^o
135^o	111^o	58^o
148^o	58^o	45^o
169^o	191^o	169^o
191^o	212^o	212^o
212^o	237^o	237^o
225^o	237^o	249^o
237^o	237^o	249^o
249^o	249^o	249^o
262^o	291^o	291^o
278^o	291^o	262^o
291^o	278^o	303^o
303^o	249^o	249^o
315^o	303^o	278^o
328^o	303^o	237^o
349^o	225^o	237^o

Actual	AAMRL HRTF	RBF HRTF
10^o	169^o	148^o
32^o	111^o	123^o
45^o	98^o	82^o
58^o	82^o	69^o
69^o	82^o	82^o
82^o	58^o	58^o
98^o	123^o	82^o
111^o	69^o	69^o
123^o	82^o	98^o
135^o	98^o	123^o
148^o	135^o	111^o
169^o	169^o	169^o
191^o	191^o	212^o
212^o	237^o	225^o
225^o	237^o	237^o
237^o	278^o	237^o
249^o	262^o	278^o
262^o	278^o	262^o
278^o	278^o	278^o
291^o	237^o	291^o
303^o	0^o	291^o
315^o	212^o	262^o
328^o	225^o	237^o
349^o	328^o	191^o

Actual	AAMRL HRTF	RBF HRTF
10^o	135^o	169^o
32^o	111^o	98^o
45^o	123^o	58^o
58^o	82^o	123^o
69^o	58^o	82^o
82^o	82^o	82^o
98^o	69^o	82^o
111^o	69^o	111^o
123^o	82^o	98^o
135^o	58^o	148^o
148^o	135^o	148^o
169^o	148^o	148^o
191^o	249^o	191^o
212^o	212^o	225^o
225^o	291^o	262^o
237^o	278^o	237^o
249^o	262^o	249^o
262^o	278^o	278^o
278^o	303^o	262^o
291^o	291^o	249^o
303^o	291^o	303^o
315^o	262^o	303^o
328^o	291^o	315^o
349^o	315^o	212^o

Actual	AAMRL HRTF	RBF HRTF
10^o	169^o	148^o
32^o	135^o	58^o
45^o	98^o	123^o
58^o	82^o	69^o
69^o	45^o	58^o
82^o	45^o	82^o
98^o	69^o	58^o
111^o	69^o	58^o
123^o	111^o	98^o
135^o	111^o	135^o
148^o	135^o	148^o
169^o	148^o	169^o
191^o	212^o	191^o
212^o	225^o	225^o
225^o	278^o	262^o
237^o	278^o	291^o
249^o	262^o	262^o
262^o	315^o	303^o
278^o	315^o	303^o
291^o	303^o	291^o
303^o	303^o	278^o
315^o	262^o	262^o
328^o	303^o	262^o
349^o	291^o	191^o

Actual	AAMRL HRTF	RBF HRTF
10^o	10^o	10^o
32^o	45^o	32^o
45^o	32^o	32^o
58^o	69^o	32^o
69^o	45^o	32^o
82^o	45^o	82^o
98^o	58^o	45^o
111^o	45^o	45^o
123^o	69^o	45^o
135^o	58^o	32^o
148^o	32^o	32^o
169^o	10^o	10^o
191^o	10^o	10^o
212^o	315^o	328^o
225^o	315^o	291^o
237^o	315^o	315^o
249^o	278^o	303^o
262^o	291^o	315^o
278^o	303^o	303^o
291^o	278^o	303^o
303^o	315^o	278^o
315^o	315^o	303^o
328^o	328^o	328^o
349^o	328^o	349^o

Actual	AAMRL HRTF	RBF HRTF
10^o	148^o	10^o
32^o	58^o	82^o
45^o	69^o	98^o
58^o	69^o	58^o
69^o	69^o	98^o
82^o	82^o	111^o
98^o	98^o	58^o
111^o	111^o	82^o
123^o	69^o	123^o
135^o	111^o	69^o
148^o	69^o	111^o
169^o	123^o	123^o
191^o	10^o	212^o
212^o	237^o	225^o
225^o	249^o	237^o
237^o	278^o	249^o
249^o	249^o	278^o
262^o	291^o	278^o
278^o	278^o	262^o
291^o	249^o	291^o
303^o	262^o	303^o
315^o	249^o	262^o
328^o	237^o	303^o
349^o	237^o	237^o

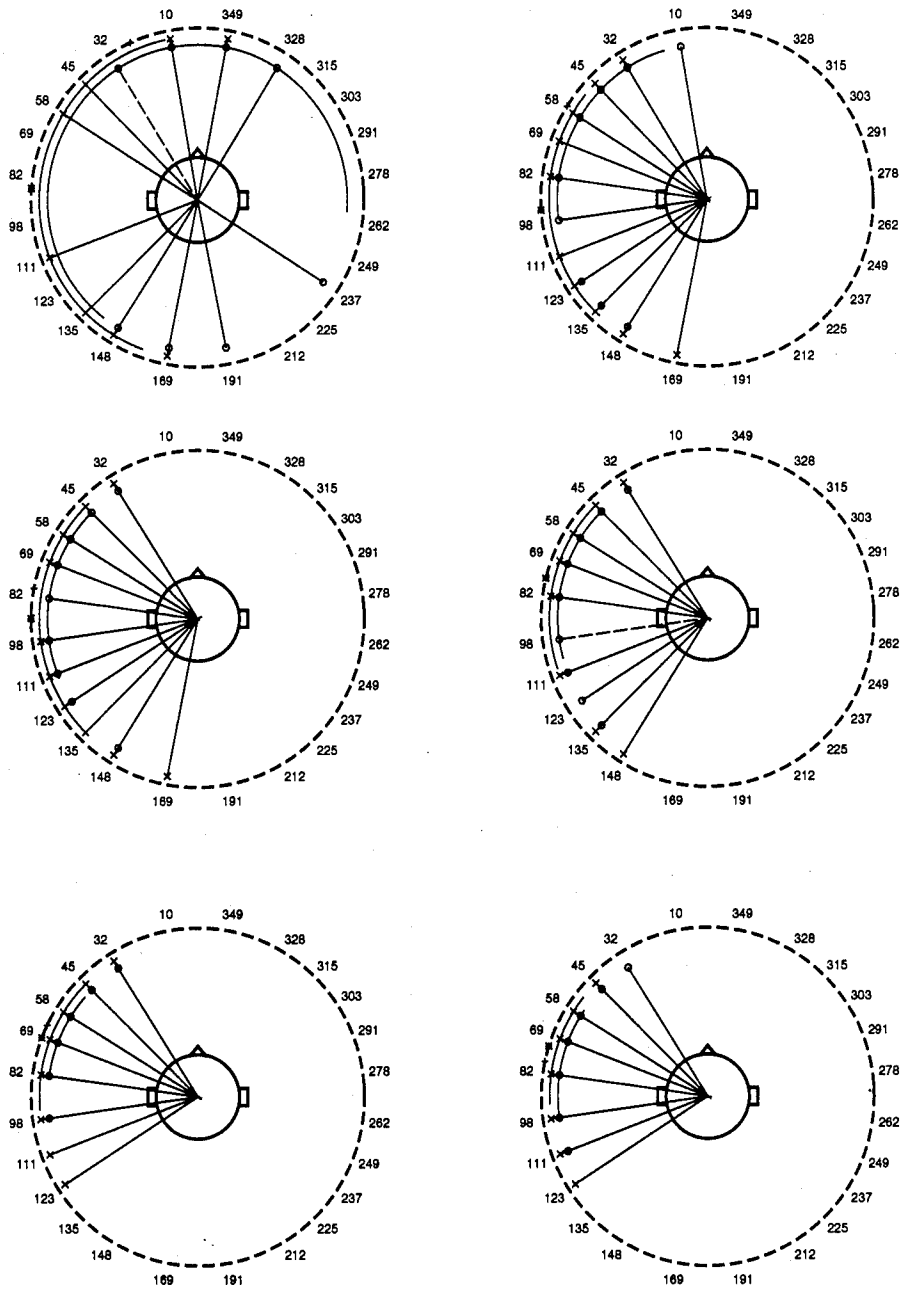


Figure 69. All subjects' responses for actual locations, 10^o, 32^o, 45^o, 58^o, 69^o and 82^o shown left to right, top to bottom respectively. Notation: $\times$ and solid line indicates with AAMRL HRTF, $\circ$ and dashed line indicate RBF HRTF. The arcs show the standard deviation of the samples. The $\star$ and $+$ indicates the mean AAMRL and RBF HRTF respectively.

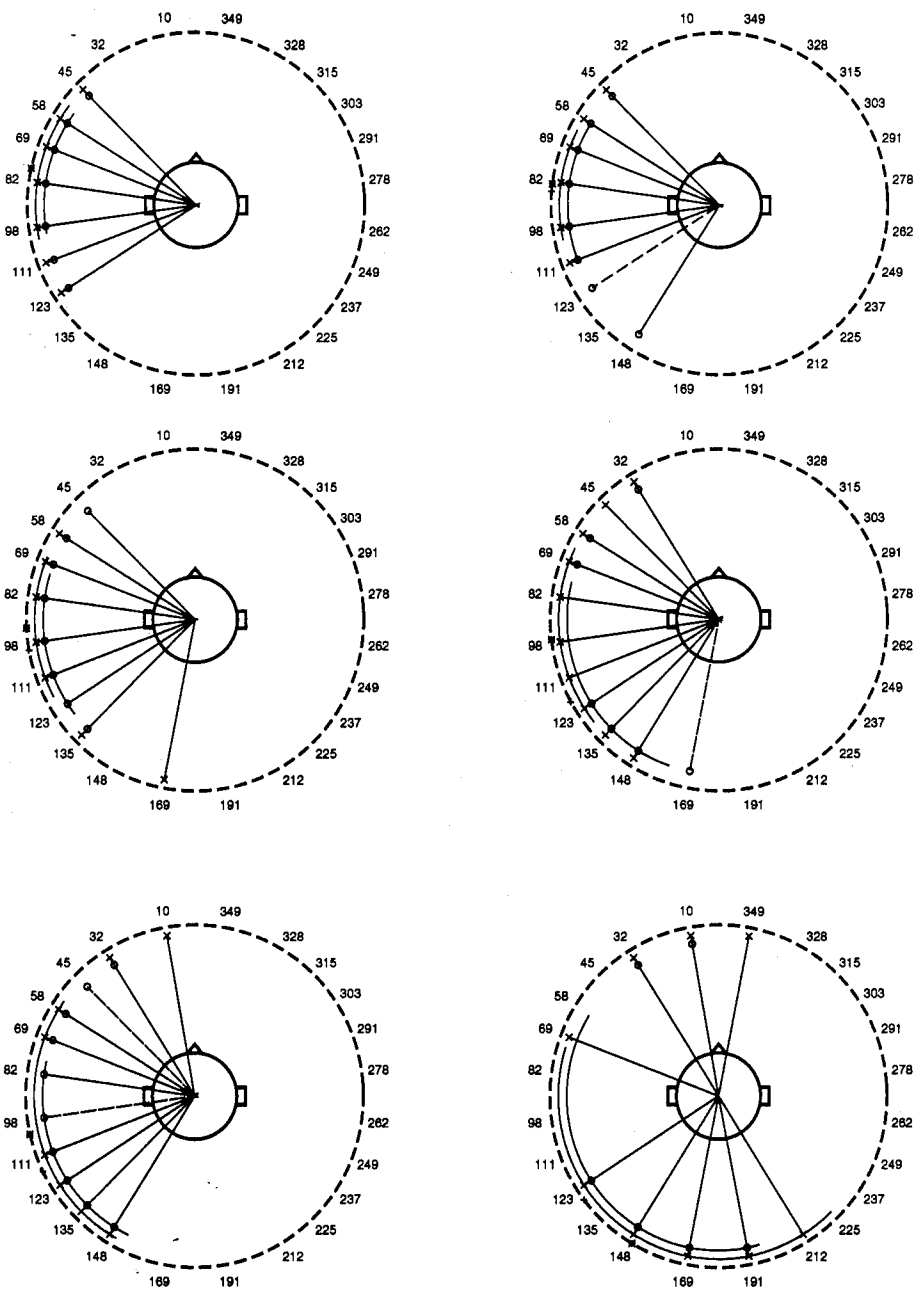


Figure 70. All subjects' responses for actual locations, 98°, 111°, 123°, 135°, 148° and 169° shown left to right, top to bottom respectively. Notation is same as above.

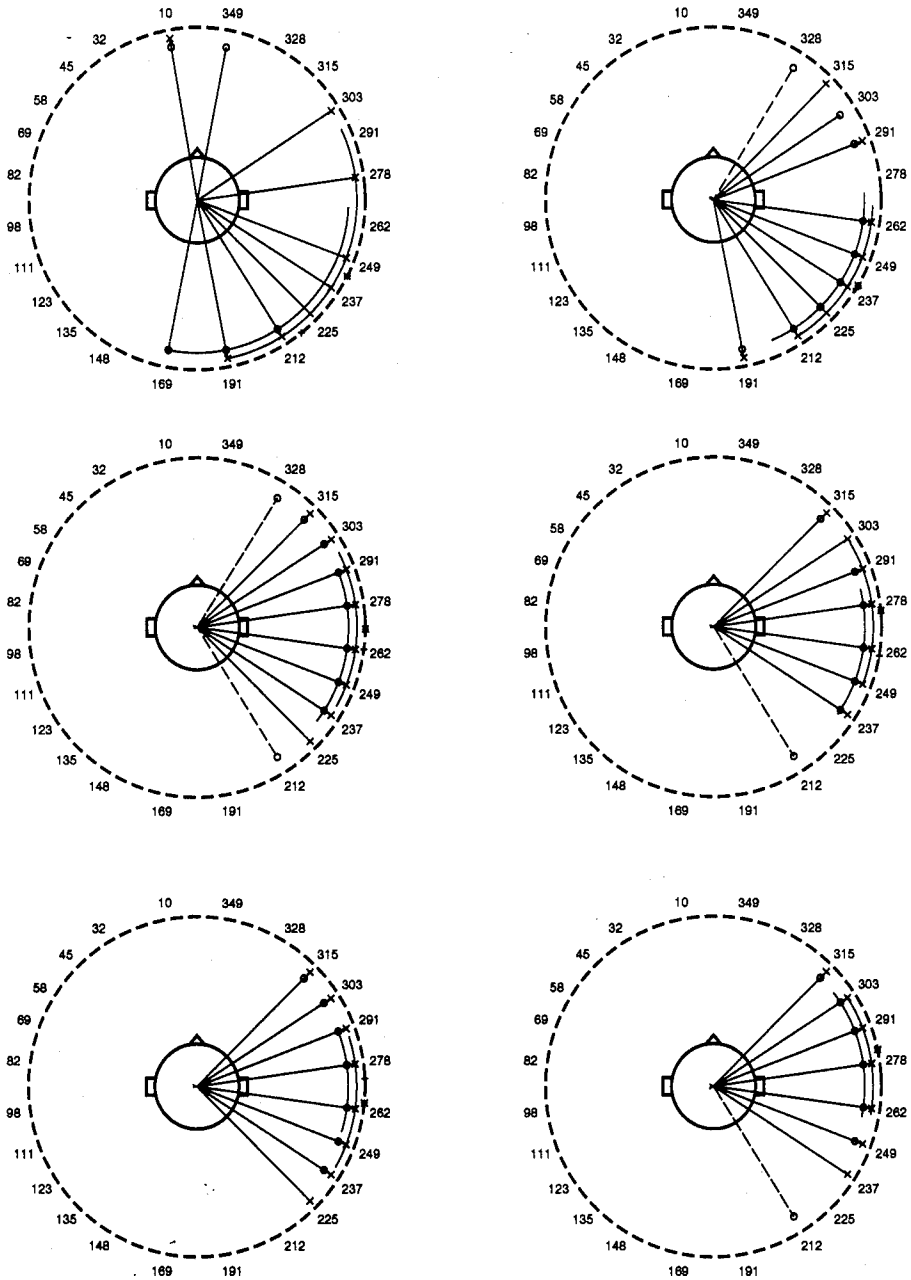


Figure 71. All subjects' responses for actual locations, 191°, 212°, 225°, 237°, 249° and 262° shown left to right, top to bottom respectively. Notation is same as above.

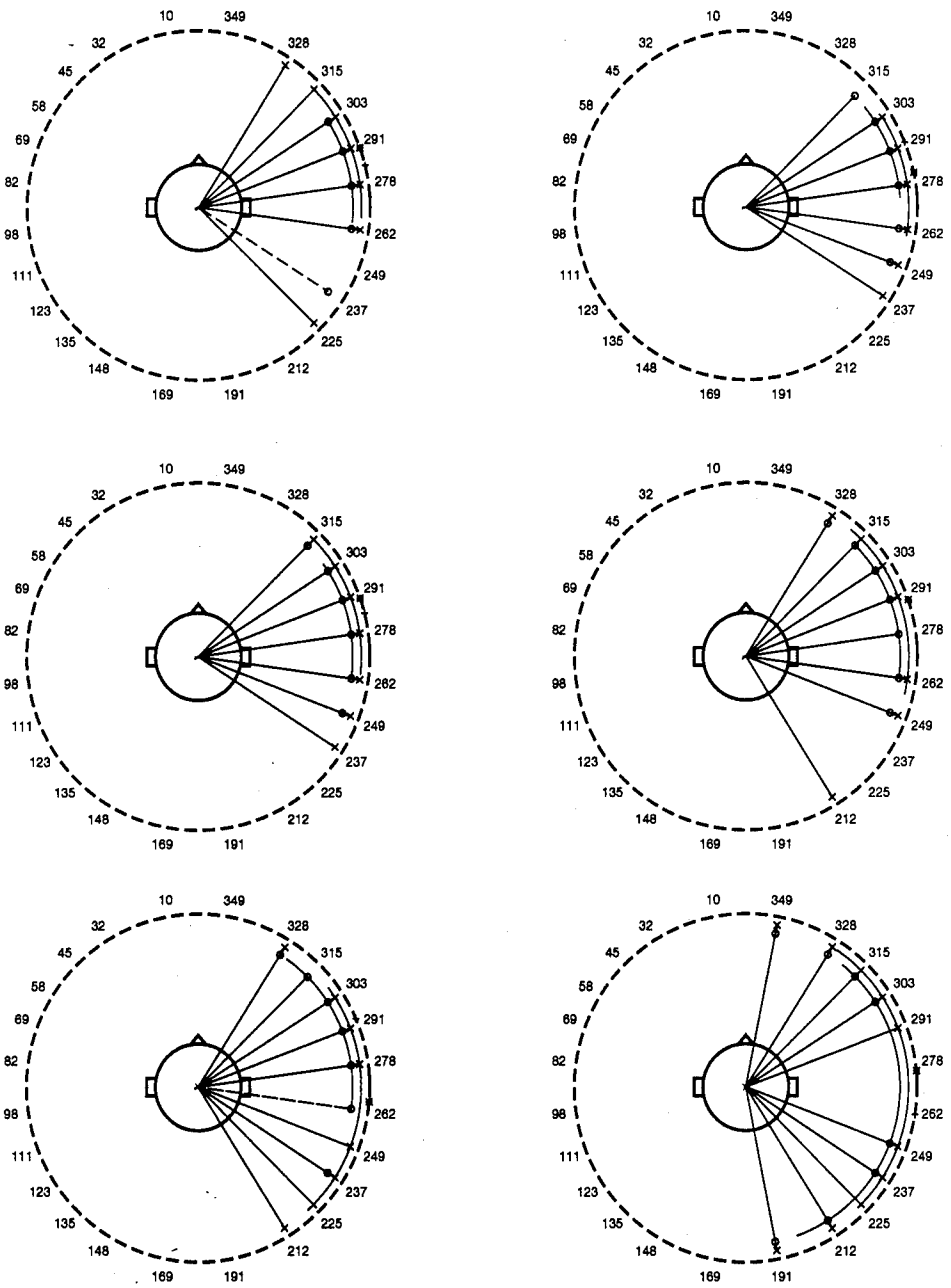


Figure 72. All subjects' responses for actual locations, 278°, 291°, 303°, 315°, 328° and 349° shown left to right, top to bottom respectively. Notation is same as above.

Appendix C. Matlab M-files

A sampling of the graphic commands in Matlab is listed in table 16.

Line and area fill commands.	
plot3	Plot lines and points in 3-D.
fill3	Draw filled 3-D polygons in 3-D.
comet3	3-D comet-like trajectories.
Contour and other 2-D plot of 3-D data.	
contour	Contour plot.
contour3	3-D contour plot.
clabel	Contour plot elevation labels.
contourc	Contour plot computation .
pcolor	Pseudocolor (checkerboard) plot.
quiver	Quiver plot.
Surface and mesh plots.	
mesh	3-D mesh surface.
meshc	Combination mesh/contour plot.
meshz	3-D Mesh with zero plane.
surf	3-D shaded surface.
surfc	Combination surf/contour plot.
surfl	3-D shaded surface with lighting.
waterfall	Waterfall plot.

3-D objects.	
cylinder	Generate cylinder.
sphere	Generate sphere.
Volume visualization.	
slice	Volumetric visualization plots.
Graph appearance.	
view	3-D graph viewpoint specification.
viewmtx	View transformation matrices.
hidden	Mesh hidden line removal mode.
shading	Color shading mode.
axis	Axis scaling and appearance.
caxis	Pseudocolor axis scaling.
colormap	Color look-up table.
Graph annotation.	
text	Text annotation.
title	Graph title.
xlabel	X-axis label.
ylabel	Y-axis label.
zlabel	Z-axis label for 3-D plots.

Table 16. Three dimensional graphic commands. Copyright (c) 1984-93 by The MathWorks, Inc.

C.1 angleresponse.m

```
function [meanerr, stdall]=angleresponse(allsubjects);
% ANGLERESPONSE Determines the mean error and standard deviation from file
% allsubjects for each angle of azimuth in experiments one and two.
% Written by Capt John K. Millhouse

[m,n]=size(allsubjects);
% correct for bad responses
for i=1:m,
    for j=1:n,
        if allsubjects(i,j)==0,
           allsubjects(i,j)=allsubjects(i,1);
        end
        if abs(allsubjects(i,j)-allsubjects(i,1))<180,
```

```
            errall(i,j)=allsubjects(i,j)-allsubjects(i,1);
        else
            errall(i,j)=allsubjects(i,j)-allsubjects(i,1)-360;
        end
    end
end
az=errall(1:24,1);
% calculate mean and std and plot results
for i=1:24,
    meanerr(i,1:3)=mean(errall(i:24:480,1:3));
    stdall(i,1:3)=(std(errall(i:24:480,1:3)));
    theta1=zeros(40,1);
    theta1(1:2:40)=allsubjects(i:24:480,2)*pi/180;
    rho1=zeros(40,1);
    rho1(1:2:40)=ones(20,1);
    theta2=zeros(40,1);
    theta2(1:2:40)=allsubjects(i:24:480,3)*pi/180;
    rho2=zeros(40,1);
    rho2(1:2:40)=.95*ones(20,1);
    mypolar(theta1,rho1,'g-')
    hold on
    mypolar(theta2,rho2,'r:')
    mypolar(allsubjects(i:24:480,2)*pi/180,ones(20,1),'gx')
    mypolar(allsubjects(i:24:480,3)*pi/180,.95*ones(20,1),'ro')
    mypolar([0 (allsubjects(i,1)+meanerr(i,2))*pi/180],...
            [0 1.05],'g*')
    mypolar([0 (allsubjects(i,1)+meanerr(i,3))*pi/180],...
            [0 1.05],'r*')
    hold off
    title(['azimuth = ',...
        num2str(allsubjects(i,1)),', ',...
        num2str(meanerr(i,2)),', ',...
        num2str(meanerr(i,3))])
    pause
end
```

C.2 anova.m

```
% get data %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
allsubjects=allsubjects2;
trial=1;
index=0;
actual=allsubjects(:,1);
with=allsubjects(:,2);
without=allsubjects(:,3);
```

```
% correct for bad responses...make mean of other subjects
%without(107)=122.78;
%without(275)=122.78;
la=length(actual);
for i=1:la
    index=index+1;
    if index==25, index=1; end
    if with(i)==0, with(i)=sum(with(index:24:480))/19;
    end
    if without(i)==0, without(i)=sum(without(index:24:480))/19;
    end
end
for i=1:la
    if with(i)<180,
       rwith(i)=180-with(i);
    else
       rwith(i)=540-with(i);
    end
    if without(i)<180,
       rwithout(i)=180-without(i);
    else
       rwithout(i)=540-without(i);
    end
end


witherr=zeros(480,1);
withouterr=zeros(480,1);
rwitherr=zeros(480,1);
rwithouterr=zeros(480,1);
for i=1:la,
    witherr(i)=actual(i)-with(i);
    withouterr(i)=actual(i)-without(i);
    rwitherr(i)=actual(i)-rwith(i);
    rwithouterr(i)=actual(i)-rwithout(i);
    if abs(witherr(i))>180,
       witherr(i)=abs(360 - abs(witherr(i)));
    else
       witherr(i)=abs(witherr(i));
    end
    if abs(withouterr(i))>180,
       withouterr(i)=abs(360 - abs(withouterr(i)));
    else
       withouterr(i)=abs(withouterr(i));
    end
```

```
    if abs(rwitherr(i))>180,
       rwitherr(i)=abs(360 - abs(rwitherr(i)));
    else
       rwitherr(i)=abs(rwitherr(i));
    end
    if abs(rwithouterr(i))>180,
       rwithouterr(i)=abs(360 - abs(rwithouterr(i)));
    else
       rwithouterr(i)=abs(rwithouterr(i));
    end
end
%correct for reversal
revcount=0;
for i=1:la,
    if witherr(i)>rwitherr(i),
       witherr(i)=rwitherr(i);
       revcount=revcount+1;
    end
    if withouterr(i)>rwithouterr(i),
       withouterr(i)=rwithouterr(i);
       revcount=revcount+1;
    end
end

% start anova %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
X=[witherr withouterr];
a=2;
b=24;
c=20;
n=1;
% get totals %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Ti=zeros(a,1);
Tj=zeros(b,1);
Tk=zeros(c,1);
Tij=zeros(a,b);
Tij2=zeros(a,b);
Tik=zeros(a,c);
Tjk=zeros(b,c);
for i=1:a
for j=1:b
for k=1:c
    Ti(i)=Ti(i)+getx(X,i,j,k);
    Tj(j)=Tj(j)+getx(X,i,j,k);
    Tk(k)=Tk(k)+getx(X,i,j,k);
    Tij(i,j)=Tij(i,j)+getx(X,i,j,k);
```

```
    Tik(i,k)=Tik(i,k)+getx(X,i,j,k);
    Tjk(j,k)=Tjk(j,k)+getx(X,i,j,k);
end
end
end
Tijk=X;
T=sum(sum(X));
% get sum of squares %%%%%%%%%%%%%%%%%%%%%%
SSa=0; SSb=0; SSc=0;
SSab=0; SSac=0; SSbc=0;
SSt=0; SSe=0;
for i=1:a
    SSa=SSa+ ((Ti(i).^2)/(b*c));
end
for j=1:b
    SSb=SSb+ Tj(j).^2/(a*c);
end
for k=1:c
    SSc=SSc+ Tk(k).^2/(a*b);
end
for i=1:a
for j=1:b
    SSab=SSab+ Tij(i,j).^2/(c);
end
end
for i=1:a
for k=1:c
    SSac=SSac+ Tik(i,k).^2/(b);
end
end
for j=1:b
for k=1:c
    SSbc=SSbc+ Tjk(j,k).^2/(a);
end
end
SSab=SSab - SSa - SSb + T.^2/(a*b*c);
SSac=SSac - SSa - SSc + T.^2/(a*b*c);
SSbc=SSbc - SSb - SSc + T.^2/(a*b*c);
SSa=SSa- T.^2/(a*b*c);
SSb=SSb- T.^2/(a*b*c);
SSc=SSc- T.^2/(a*b*c);
for i=1:a
for j=1:b
for k=1:c
    SSt=SSt+ getx(X,i,j,k).^2;
```

```
end
end
end
SSt=SSt- T.^2/(a*b*c);
SSe= SSt - SSa - SSb - SSc - SSab - SSac - SSbc;

% mean square
MSa=SSa/(a-1);
MSb=SSb/(b-1);
MSc=SSc/(c-1);
MSab=SSab/((a-1)*(b-1));
MSac=SSac/((a-1)*(c-1));
MSbc=SSbc/((b-1)*(c-1));
DFe=(a*b*c - 1)-(a-1)-(b-1)-(c-1)-((a-1)*(b-1))-((a-1)*(c-1))-((b-1)*(c-1));
MSe=SSe/DFe;

% F value
Fa=MSa/MSe;
Fb=MSb/MSe;
Fc=MSc/MSe;
Fab=MSab/MSe;
Fac=MSac/MSe;
Fbc=MSbc/MSe;

% display
table= [1  a-1          SSa  MSa  Fa
        2  b-1          SSb  MSb  Fb
        3  c-1          SSc  MSc  Fc
        12 (a-1)*(b-1)  SSab MSab Fab
        13 (a-1)*(c-1)  SSac MSac Fac
        23 (b-1)*(c-1)  SSbc MSbc Fbc
        0  DFe          SSe  MSe   0
        0  (a*b*c -1)   SSt   0    0]
```

C.3 drc.m

```
function [range,minvalue,maxvalue]=drc(matrix);

% Dynamic Range Compression (drc)
% return range between 0 and 1 of input matrix columns
% Written by Capt John K. Millhouse

[m n]=size(matrix);

temp=matrix-(ones(m,1)*min(matrix));
```

```
range=temp./(ones(m,1)*max(temp));
minvalue=min(matrix);
maxvalue=max(temp);
```

C.4 expmat1.m

```
% Setup for experiment one
% Creates a script file called 'run_experiment'
% Written by Capt John K. Millhouse

expaz=[10 32 45 58 69 82 98 111 123 135 148 169 191 212 225 237
249 262 278 291 303 315 328 349 1000 3200 4500 5800 6900 8200
9800 11100 12300 13500 14800 16900 19100 21200 22500 23700 24900
26200 27800 29100 30300 31500 32800 34900];

% randomly place sounds
row=randperm(48);

for i=1:48,
    snd(i)=expaz(row(i));
    temp=['0' int2str(snd(i))];
    if snd(i)< 100
       tsnd(i,1:3)=temp(1:3);
    elseif snd(i) < 1000
       tsnd(i,1:3)=temp(2:4);
    elseif snd(i) < 10000
       tsnd(i,1:3)=temp(1:3);
    else
       tsnd(i,1:3)=temp(2:4);
    end
end

fid = fopen('run_experiment','w');
fprintf(fid,'#!/bin/csh %s\n',' ');
fprintf(fid,'makecompass &%s\n',' ');
fprintf(fid,'introduction %s\n',' ');
fprintf(fid,'set ans = $< %s\n',' ');


for i=1:48,
  if snd(i)<1000
     tex1=['s32cplay -f40000 ../actualfilters/rl' int2str(snd(i)) '.d %s\n'];
     tex2=['s32cplay -f40000 ../actualfilters/rl' int2str(snd(i)) '.d ;
           echo ''Aactual ' tsnd(i,:) ''' >>  results.out %s\n'];
  else
```

```
    tex1=['s32cplay -f40000 ../meanHRTF/rl' int2str(snd(i)/100) '.d %s\n'];
    tex2=['s32cplay -f40000 ../meanHRTF/rl' int2str(snd(i)/100) '.d ;
          echo ''Bactual ' tsnd(i,:) ''' >>  results.out %s\n'];
  end
    fprintf(fid,tex1,' ');
    fprintf(fid,'sleep 1 %s\n',' ');
    fprintf(fid,tex2,' ');
    fprintf(fid,'clear %s\n',' ');
    fprintf(fid,'echo ''Select location and press return'' %s\n',' ');
    fprintf(fid,'set ans = $< %s\n',' ');
  end
end

fprintf(fid,'s32cplay -f40000 thankyou.d %s\n',' ');

fclose(fid);
```

C.5 expmat2.m

```
% Setup for experiment two
% Creates a script file called 'run_experiment'
% Written by Capt John K. Millhouse

expaz=[10 32 45 58 69 82 98 111 123 135 148 169 191 212 225 237
249 262 278 291 303 315 328 349 1000 3200 4500 5800 6900 8200
9800 11100 12300 13500 14800 16900 19100 21200 22500 23700 24900
26200 27800 29100 30300 31500 32800 34900];

% randomly place sounds
row=randperm(48);

for i=1:48,
    snd(i)=expaz(row(i));
    temp=['0' int2str(snd(i))];
    if snd(i)< 100
       tsnd(i,1:3)=temp(1:3);
    elseif snd(i) < 1000
       tsnd(i,1:3)=temp(2:4);
    elseif snd(i) < 10000
       tsnd(i,1:3)=temp(1:3);
    else
       tsnd(i,1:3)=temp(2:4);
    end
end
```

```
fid = fopen('run_experiment','w');
fprintf(fid,'#!/bin/csh %s\n',' ');
fprintf(fid,'makecompass &%s\n',' ');
fprintf(fid,'introduction %s\n',' ');
fprintf(fid,'set ans = $< %s\n',' ');


for i=1:48,
  if snd(i)<1000
     tex1=['echo ''Aactual ' tsnd(i,:) ''' >>  results.out;
            s32cplay -f40000 ../actualfilters/rl' int2str(snd(i)) '.d %s\n'];
     tex2=['s32cplay -f40000 ../actualfilters/rl' int2str(snd(i)) '.d %s\n'];
  else
     tex1=['echo ''Bactual ' tsnd(i,:) ''' >>  results.out;
            s32cplay -f40000 ../netfilters/rl' int2str(snd(i)/100) '.d %s\n'];
     tex2=['s32cplay -f40000 ../netfilters/rl' int2str(snd(i)/100) '.d %s\n'];
  end
    fprintf(fid,tex1,' ');
    fprintf(fid,'sleep 1 %s\n',' ');
    fprintf(fid,tex2,' ');
    fprintf(fid,'clear %s\n',' ');
    fprintf(fid,'echo ''Select location and press return'' %s\n',' ');
    fprintf(fid,'set ans = $< %s\n',' ');
  end
end

fprintf(fid,'s32cplay -f40000 thankyou.d %s\n',' ');

fclose(fid);
```

C.6 idrc.m

```
function [matrix]=idrc(range,minvalue,maxvalue);

% inverse Dynamic Range Compression (idrc)
% return matrix to original values
% Written by Capt John K. Millhouse

[m n]=size(range);

temp=range.*(ones(m,1)*maxvalue);
matrix=temp+(ones(m,1)*minvalue);
```

C.7 LRsphere.m

```
% Create left and right spheres of HRTF data
% Written by Capt John K. Millhouse
k=1;
for i=1:2:90,
freq=i;
[x y z c]=makesphere(features,freq,16);

subplot(1,2,1);
surf(x,y,z,c);
caxis([0, 0.7915]);
view(180,0);
shading interp
axis off
title(num2str(features(i,4)))
text(.1,1,0,'(&) <-- Right ear')
xlabel('right side of head');

subplot(1,2,2);
surf(x,y,z,c);
caxis([0, 0.7915]);
view(0,0);
shading interp
axis off
text(-.1,-1,0,'(&) <-- Left ear')
text(-1.8,0,0,'<-- Nose -->')
xlabel('left side of head');
title('Hz');

colormap(jet)
k=k+1;
end
```

C.8 makesphere.m

```
function [x ,y ,z, c]= makesphere(features,freq,size)
% freq range 1 to 93
% Create sphere with color corresponding to HRTF value
% Written by Capt John K. Millhouse
k=0;
clear az elev resp
for i=freq:93:25296,
  k=k+1;
  az(k)=features(i,2);
  elev(k)=features(i,3);
```

```
  resp(k)=features(i,5);
end
c=mapsurf(az,elev,resp,size);
[x y z]=sphere(size);
```

C.9 mapsurf.m

```
function [c]  = mapsurf(az,elev,value,spheresize)
% maps the HRTF value to the az and elev
% locations on a sphere
% Written by Capt John K. Millhouse

n=length(az);
I=0;

c = zeros(spheresize+1,spheresize+1);
v = c;

for I = 1:n,
   a = fix(az(I)/360*(spheresize+1))+1;
   e = spheresize+1 - fix((elev(I)+90)/180*(spheresize+1));
   if e>spheresize+1
      e=spheresize+1;
   end
   if e<1
      e=1;
   end
   if a>spheresize+1
      a=spheresize+1;
   end
   if a<1
      a=1;
   end
   v(e,a)=v(e,a)+1;
   c(e,a)=c(e,a)+value(I);
end

for I = 1:spheresize+1,
    for J = 1:spheresize+1,
       if v(I,J)==0
       else
         c(I,J) = c(I,J)/v(I,J);
       end
    end
```

```
end
```

C.10 mapeval.m

```
function [actual, netout, error]=mapeval(x, y, data);
% Written by Capt John K. Millhouse
% Converts LNKmap output to Matlab matrix suitable for
% display
% data in form [number actual netout error]

nx=length(x);
ny=length(y);
actual=zeros(nx,ny);
netout=zeros(nx,ny);
error=zeros(nx,ny);
k=0;
for i=1:nx,
for j=1:ny,
     k=k+1;
     actual(i,j)=data(k,2);
     netout(i,j)=data(k,3);
     error(i,j)=data(k,4);
end
end
actual=actual';
netout=netout';
error=error';
```

C.11 mypolar.m

```
function hpol = mypolar(theta,rho,line_style)
%POLAR  Polar coordinate plot.
%       POLAR(THETA, RHO) makes a plot using polar coordinates of
%       the angle THETA, in radians, versus the radius RHO.
%       POLAR(THETA,RHO,S) uses the linestyle specified in string S.
%       See PLOT for a description of legal linestyles.
%
%       See also PLOT, LOGLOG, SEMILOGX, SEMILOGY.

%       Copyright (c) 1984-93 by The MathWorks, Inc.
%       Modified by Capt John K. Millhouse to display
%       head figure and test locations.

if nargin < 1
        error('Requires 2 or 3 input arguments.')
elseif nargin == 2
```

```
        if isstr(rho)
                line_style = rho;
                rho = theta;
                [mr,nr] = size(rho);
                if mr == 1
                        theta = 1:nr;
                else
                        th = (1:mr)';
                        theta = th(:,ones(1,nr));
                end
        else
                line_style = 'auto';
        end
elseif nargin == 1
        line_style = 'auto';
        rho = theta;
        [mr,nr] = size(rho);
        if mr == 1
                theta = 1:nr;
        else
                th = (1:mr)';
                theta = th(:,ones(1,nr));
        end
end
if isstr(theta) | isstr(rho)
        error('Input arguments must be numeric.');
end
if any(size(theta) ~= size(rho))
        error('THETA and RHO must be the same size.');
end

% get hold state
cax = newplot;
next = lower(get(cax,'NextPlot'));
hold_state = ishold;

% get x-axis text color so grid is in same color
tc = get(cax,'xcolor');

% only do grids if hold is off
if ~hold_state

% make a radial grid
        hold on;
        hhh=plot([0 max(theta(:))],[0 max(abs(rho(:)))]);
```

```
        v = [get(cax,'xlim') get(cax,'ylim')];
        ticks = length(get(cax,'ytick'));
        delete(hhh);
% check radial limits and ticks
        rmin = 0; rmax = v(4); rticks = ticks-1;
        if rticks > 5   % see if we can reduce the number
                if rem(rticks,2) == 0
                        rticks = rticks/2;
                elseif rem(rticks,3) == 0
                        rticks = rticks/3;
                end
        end

% define a circle
        th = 0:pi/50:2*pi;
        xunit = cos(th);
        yunit = sin(th);
% now really force points on x/y axes to lie on them exactly
    inds = [1:(length(th)-1)/4:length(th)];
    xunits(inds(2:2:4)) = zeros(2,1);
    yunits(inds(1:2:5)) = zeros(3,1);
    ynose=[.24; .3; .24]*rmax;
    xnose=[.05; 0; -.05]*rmax;
    year=[.05; .05; -.05; -.05]*rmax;
    xear=[.24; .3; .3; .24]*rmax;

        rinc = (rmax-rmin)/rticks;
        i=rmax;
        plot(xunit*i,yunit*i,':','color',tc,'linewidth',1);
        i=rmax/4;
        plot(xunit*i,yunit*i,'-','color',tc,'linewidth',1);
        plot(xnose,ynose,'-','color',tc,'linewidth',1);
        plot(xear,year,'-','color',tc,'linewidth',1);
        plot(-xear,year,'-','color',tc,'linewidth',1);



% plot spokes
        th = [10 32 45 58 69 82 98 111 123 135 148 169 191
              212 225 237 249 262 278 291 303 315 328 349]*pi/180;
        cst = -sin(th);
        snt = cos(th);
        cs = [zeros(1,24); cst];
        sn = [zeros(1,24); snt];
        %plot(rmax*cs,rmax*sn,':','color',tc,'linewidth',1);
```

```
% annotate spokes in degrees
        rt = 1.1*rmax;
        for i = 1:max(size(th))
            text(rt*cst(i),rt*snt(i),int2str(th(i)*180/pi),
                 'horizontalalignment','center')
        end

% set viewto 2-D
        view(0,90);
% set axis limits
        axis(rmax*[-1 1 -1.1 1.1]);
end

% transform data to Cartesian coordinates.
xx = rho.*(-sin(theta));
yy = rho.*cos(theta);

% plot data on top of grid
if strcmp(line_style,'auto')
        q = plot(xx,yy);
else
        q = plot(xx,yy,line_style);
end

if nargout > 0
        hpol = q;
end
if ~hold_state
        axis('equal');axis('off');
end

% reset hold state
if ~hold_state, set(cax,'NextPlot',next); end
```

C.12 plotresult.m

```
function [h]=plotresult(result,rquestion);
% plot results of a subjects response
% Written by Capt John K. Millhouse
if nargin < 2
   rquestion = 'nor';
end
rhos10=ones(1,360);
rhos20=ones(1,360);
```

```
for i=1:length(result),
 theta1=result(i,2)*pi/180;
 theta2=(result(i,3))*pi/180;
 lt='yo';
 if rquestion=='rev',
    if abs(result(i,4))>abs(result(i,6))
       theta2=(result(i,5))*pi/180;
       lt='r+';
    end
 end
 index=round(theta2*180/pi)+1;
 if result(i,1)==10
    rhos10(index)=rhos10(index)-.05;
    rho10=rhos10(index);
    subplot(1,2,1)
    mypolar(theta1,1,'gx')
    hold on
    mypolar(theta2,rho10,lt)
    mypolar([theta1; theta2],[1; rho10],'g-')
    title('With HRTF')
 else
    rhos20(index)=rhos20(index)-.05;
    rho20=rhos20(index);
    subplot(1,2,2)
    mypolar(theta1,1,'mx')
    hold on
    mypolar(theta2,rho20,lt)
    mypolar([theta1; theta2],[1; rho20],'m-')
    title('Without HRTF')
 end
end

for i=1:2
subplot(1,2,i)
hold off
end
```

C.13 plothist.m

```
function [correct10theta,correct20theta]=
         plothist(testall,test10,test20,rquestion,correctdeg);
% Written by Capt John K. Millhouse
% Plots Polar histogram of data of correct responses
% testall is all test responses
```

```
% test10 is test responses for filter1
% test20 is test responses for filter2
% rquestion is 'rev' reverses or 'nor' no reverses
% correctdeg is the number of degrees off an still correct

if nargin<5
   correctdeg=23;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% plot histograms
actualtheta=testall(:,2)*pi/180;
testtheta=testall(:,3)*pi/180;
testwrevtheta=[test10(:,3)*pi/180; test20(:,3)*pi/180];
test10revtheta=[test10(:,3)*pi/180];
test20revtheta=[test20(:,3)*pi/180];
test10theta=testall(1:length(test10),3)*pi/180;
test20theta=testall(length(test10)+1:length(testall),3)*pi/180;
for i=1:length(test10),
   if abs(test10(i,4))<correctdeg,
      correct10theta=[correct10theta; test10(i,2)*pi/180];
   end
end
for i=1:length(test20),
   if abs(test20(i,4))<correctdeg,
      correct20theta=[correct20theta; test20(i,2)*pi/180];
   end
end
bins=180;
% number of bins in the histogram
% if the number of bins is >90 then offset the bins for display
if bins>90
   correct10theta=correct10theta+(pi/90);
   correct20theta=correct20theta-(pi/90);
   test10theta=test10theta+(pi/90);
   test20theta=test20theta-(pi/90);
end
[tact,ract]=rose(actualtheta,bins);
[tc10,rc10]=rose(correct10theta,bins);
[tc20,rc20]=rose(correct20theta,bins);
[t10,r10]=rose(test10theta,bins);
[t20,r20]=rose(test20theta,bins);

figure
mypolar(tact,ract/2,'w:')
hold on
mypolar(t10,r10,'w-.')
```

```
mypolar(t20,r20,'w-')
title(['Histogram of all responses'])
legend('w:','Actual','w-.','with HRTF','w-','without HRTF')
hold off
figure
mypolar(tact,ract/2,'w:')
hold on
mypolar(tc10,rc10,'w-.')
mypolar(tc20,rc20,'w-')
title(['Histogram of correct responses ' '(' rquestion ')'])
legend('w:','Actual','w-.','with HRTF correct','w-','without HRTF correct')

hold off
```

C.14 readall.m

```
function [rmserr,stderr,testall,revall,test10,test20]=
         readall(rquestion,subjects);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Read all the test subjects data and save to a   %
% file named allsubjects                          %
% by J. Millhouse                                 %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%folder=['meanHRTFresults'];
folder=['results'];

if nargin < 1
   rquestion = 'nor';
end
if nargin<2
   if length(folder)==15
      % meanHRTFresults %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
      subjects=['bsmith1.out','bsmith2.out','jmillhouse1.out',...
                'jmillhouse2.out','bmcq1.out','bmcq2.out',...
                'ceisenbies1.out','ceisenbies2.out','gharrup1.out',...
                'lmyers1.out','lmyers2.out','rdennery1.out',...
                'sstewart1.out','twilson1.out','twilson2.out',...
                'amillhouse1.out','bob1.out','jrmillhouse1.out',...
                'djennings1.out','rjreid1.out'];
   else
      % results %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
      subjects=['ksmith1.out','rdennery1.out','jmillhouse1.out',...
```

```
                'jmillhouse2.out','bsmith1.out','twilson1.out',...
                'gharrup1.out','jrmillhouse1.out','jrmillhouse2.out',...
                'lmyers2.out','lmyers3.out','ddouglas1.out',...
                'ddouglas2.out','rmillhouse1.out','rmillhouse2.out',...
                'ceisenbies1.out','ceisenbies2.out','djennings1.out',...
                'srogers1.out','amillhouse1.out'];
   end

end
testall=[];
testall2=[];
revall=[];
start=1;
finish=1;
for i=4:length(subjects)
  j=i-3;
  if subjects(j:i)=='.out'
     finish=i;
     name=subjects(start:finish-4)
     [test rev]=readresult([folder '/' name '.out']);
     %figure
     %plotresult(test,rquestion);
     start=finish+1;
     testall=[testall; test];
     revall=[revall; rev];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% make output format
eval([name '=zeros(24,3);']);
f=sortdata(test(:,1:3),2);
c=0;
oldaz=0;
for a=1:length(f)
    if f(a,1)==10,
       b=2;
    else
       b=3;
    end
    az=f(a,2);
    if oldaz<az c=c+1; end
    oldaz=az;
    eval([name '(c,1)=f(a,2);']);
    eval([name '(c,b)=f(a,3);']);
end
eval(['testall2=[testall2;  ' name '];']);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
  end
```

```
end
eval(['save ' folder '/allsubjects testall2 -ascii -tabs'])
testall=sortdata(testall,1);
revall=sum(revall);

test10=[];
test20=[];
test=testall;
for i=1:length(testall)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% get reversals
  if rquestion=='rev',
  if abs(testall(i,4))>abs(testall(i,6))
     test(i,4)=test(i,6);
     test(i,6)=test(i,3);
     test(i,3)=test(i,5);
     test(i,5)=test(i,6);
     test(i,6)=999;
  end
  end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% split into two
  if testall(i,1)==10
     test10=[test10; test(i,:)];
  else
     test20=[test20; test(i,:)];
  end
end

avgerr=[mean(test10(:,4)) mean(test20(:,4))];
rmserr=sqrt(avgerr.^2);
stderr=[std(test10(:,4)) std(test20(:,4))];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% plot histograms
%plothist(testall,test10,test20,rquestion);
```

C.15 removerev.m

```
function [y,z]=removerev(x);
% Written by Capt John K. Millhouse

[m,n]=size(x);
y=[];
z=[];
for i=1:m,
    if x(i,n)<999,
       y=[y; x(i,:)];
    else
```

```
      z=[z; x(i,:)];
    end
end
```

C.16 sortdata.m

```
function [ Y ] = sortdata(file1, column)

[n,m]=size(file1);
[X,I]=sort(file1);
for j=1:n,
  for k=1:m,
     if k==column
        Y(j,k)=X(j,column);
     else
        Y(j,k)=file1(I(j,column),k);
     end
  end
end
```

Bibliography

1. Abbagnaro, Louis A., et al. "Measurements of Diffraction and Interaural Delay of a Progressive Sound Wave Caused by the Human Head," *Journal of the Acoustical Society of America*, *58*:693–700 (September 1975).
2. Begault, Durand R. "Head-Up Auditory Display for Traffic Collision Avoidance System Advisories: A Preliminary Investigation," *The Journal of the Human Factors and Ergonomics society*, *35*(4):707–717 (December 1993).
3. Begault, Durand R. and Elizabeth M Wenzel. "Headphone Localization of Speech," *The Journal of the Human Factors and Ergonomics society*, *35*(2):361–376 (June 1993).
4. Blattner, Mecra M., et al. "Earcons and Icons: Their Structure and Common Design Principles," *Human-Computer Interaction*, *4*:11–44 (1989).
5. Blauert, Jens. *Spatial Hearing*. MIT Press, 1983.
6. Bly, Sara. *Sound and Computer Information Presentation*. Thesis, Computing Science Group University of California, Davis, Davis, CA, March 1982.
7. Fitts, P. *Psychological research on equipment design*. A study of location discrimination ability 19, Columbus, Ohio: Ohio State University: Army Air Force, Aviation Psychology Program., 1947.
8. Furness, T. "The Super Cockpit and its human factors challenges." *Proceedings of the Human Factors Society 30th Annual Meeting*. 48–52. Santa Monica, CA: Human Factors Society, 1986.
9. Gonzalez, Rafael C. and Richard E. Woods. *Digital Image Processing*. Addison-Wesley, 1992.
10. Kahn, M. J. "Human awakening and subsequent identification of fire-related cues." *Proceedings of the Human Factors Society 27th Annual Meeting*. 806–810. Santa Monica, CA: Human Factors Society, 1983.
11. Kuhn, George F. "Model for the Interaural Time Differences in the Azimuthal Plane," *Journal of the Acoustical Society of America*, *62*:157–167 (July 1977).
12. McKinley, Richard L. *Concept and Design of an Auditory Localization Cue Synthesizer*. MS thesis, Air Force Institute of Technology (AU), 1988.
13. Middlebrooks, John C. "Directional Sensitivity of Sound-Pressure Levels in the Human Ear Canal," *Journal of the Acoustical Society of America*, *86*:89–108 (July 1989).
14. Milton, J. S. and Jesse C. Arnold. *Introduction to Probability and Statistics* (second Edition). McGraw-Hill, 1990.
15. Moller, Henrik. "Fundamentals of Binaural Technology," *Applied Acoustics*, *36*:171–218 (March 1992).

16. Neti, Chalapathy, et al. "Neural Network Models of Sound Localization Based on Directional Filtering by the Pinna," *Journal of the Acoustical Society of America*, 3140–3156 (December 1992).

17. Oldfield, Simon R. and Simon P. A. Parker. "Acuity of Sound Localisation: A Topgraphy of Auditory Space. I. Normal Hearing Conditions," *Perception*, *13*:581–600 (1984).

18. Oldfield, Simon R. and Simon P. A. Parker. "Acuity of Sound Localisation: a topography of auditoy space. II. Pinna cues absent," *Perception*, *13*:601–617 (1984).

19. Park, Jooyoung and Irwin W. Sandberg. "Approximation and Radial-Basis-Function Networks," *Neural Computation*, *5*(2):305–316 (March 1993).

20. Rogers, Steven K., et al. *An Introduction to Biological and Artificial Neural Networks*. Wright-Patterson AFB OH, 45433: Air force Institute of Technology (AU), October 1990.

21. Ruck, Dennis W. From a conversation, May 1994.

22. Sanders, Mark S. and Ernest J. McCormick. *Human Factors in Engineering and Design* (seventh Edition). McGraw-Hill, 1993.

23. Scarborough, Captain Eric L. *Enhancement of Audio Localization Cue Synthesis by Adding Environmental and Visual Cues*. MS thesis, Air Force Institute of Technology (AU), 1992.

24. Shaw, Edgar A.G. "Acoustical Characteristics of the Human External Ear." *Conference on Binaural and Spatial Hearing*. 1993.

25. Smith, Brain A. *Binaural Room Simulation*. MS thesis, School of Engineering, Air force Institute of Technology (AU), Wright-Patterson AFB OH, December 1993.

26. Smith, Stuart, et al. "Stereophonic and Surface Sound Generation for Exploratory Data Analysis." *Chi '90 proceedings*. 125–132. April 1990.

27. Spiegel, Murray R. *Schaum's Outline of Theory and Problems of Statistics* (second Edition). Schaum's Outline Series, McGraw-Hill, 1994.

28. Stevens, S. S. and E. B. Newman. "The localization of actual sources of sound," *American Journal of Psychology*, *48*:297–306 (1936).

29. Watkins, A. J. "Psychoacoustical aspects of synthesized vertical locale cues," *Journal of the Acoustical Society of America*, *63*:1152–1165 (1978).

30. Wenzel, Elizabeth M. "Localization in Virtual Acoustic Displays," *Presence*, *I* (1992).

31. Wightman, Frederic L. and Doris J. Kistler. "Factors Affecting Relative Importance of Sound Localization Cues." *Conference on Binaural and Spatial Hearing*. 1993.

Captain John K. Millhouse was born on 11 October 1966 in Milwaukee, Wisconsin. He graduated from Centerville High School, Centerville Ohio in 1985 and attended Ohio University from which he received a Bachelor of Science degree in electrical engineering in June 1989. He was commissioned though the Air Force Reserve Officer Training Corp (AFROTC) at Ohio University and assign to the Range and Air Base Systems Program Office, Air Combat Maneuvering Instrumentation (ACMI) Directorate, Eglin AFB, Florida. He entered the Masters Program in the School of Engineering, Air Force Instituted of Technology, in June 1993.

He is married to Audra L. Millhouse (Baker) from Grafton, Ohio. They have one son James Jerome Millhouse.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1994	Master's Thesis

4. TITLE AND SUBTITLE
Head Related Transfer Function Approximation Using Neural Networks

5. FUNDING NUMBERS

6. AUTHOR(S)
John Kindell Millhouse

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GE/ENG/94D-21

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
None

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This thesis determines whether an artificial neural network (ANN) can approximate the Armstrong Aerospace Medical Research Laboratories (AAMRL) head related transfer functions (HRTF) data obtained from research at AAMRL during the fall of 1988. The first test determines whether HRTF lends any support in sound localization when compared to no HRTF (Interaural Time Delay only). There is a statistically significant interaction between the location of the sound and whether the HRTF or no HRTF is used. When this interaction is removed using the alternate *F*-Value, the statistics give the result of equal means for the filters and azimuth. This means that at certain angles of azimuth, the HRTF either provides no advantage at all or hinders localization capabilities. Adding in the corrections for reversals changes the results to where the means of the azimuth are not statistically equal. The reversal corrections will inherently reduce the error results. However, comparing the number of reversals indicates an advantage of using the HRTF over no HRTF. The second test determines whether HRTF and ANN lend the same amount of imformation in sound localization when compared to each other. With reversal corrections included, a statistical advantage is not indicated, and the means of the two filters are statistically equal. Also there is a statistically significant interaction between the location of the sound and whether the AAMRL HRTF or ANN HRTF is used. Comparing the number of reversals does not indicate a large difference between the AAMRL HRTF and the RBF HRTF.

14. SUBJECT TERMS	15. NUMBER OF PAGES
3D sound, HRTF, Neural Networks	155
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL